View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Qucosa - Publikationsserver der Universität Leipzig

Linear Approximation of Groups and Ultraproducts of Compact Simple Groups

Von der Fakultät für Mathematik und Informatik
der Universität Leipzig
angenommene

DISSERTATION

zur Erlangung des akademischen Grades

DOCTOR RERUM NATURALIUM
(Dr. rer. nat.)

im Fachgebiet

MATHEMATIK

vorgelegt von

Dipl.-Math. Stephan Abel Stolz
geboren am 7. April 1983 in Göttingen

Die Annahme der Dissertation wurde empfohlen von:
1. **Prof. Dr. Andreas Thom, Universität Leipzig**
2. **Prof. Dr. Nikolay Nikolov, University of Oxford**

Die Verleihung des akademischen Grades erfolgt mit Bestehen
der Verteidigung am 17.10.2013 mit dem Gesamtprädikat **summa cum laude**

Acknowledgements

I want to express my gratitude to all the people who helped me during the time of writing this thesis.

First of all I am indebted to Professor Andreas Thom for his support and his constant encouragement and advice.

I thank Xu Shuang for tolerating that in the last years I sometimes gave mathematics precedence over more important things.

The Yoseikan Dojo Leipzig for teaching me *doryoku*, *arigatougozaimashita*.

I thank Alessandro Carderi, Carolin Noack, Dominik Wulf, Felix Rehren, Hendrik Demmer, Julian Pook, Philip Dowerk and Ulrike Robra, who all helped finding both mathematical and typographical errors, correcting my use of the English language or optimizing the layout.

I am grateful to my colleagues Christoph Gamm, Danny Tobisch, Konrad Zimmermann, Marco Neumann, Matthias Waack, Peter Schlicht, Sebastian Kochinke and Tim Netzer at the Mathematics Institute in Leipzig for interesting discussions, helping with computer issues, finding books and much more.

Contents

1 INTRODUCTION

Since the introduction of *sofic groups* by Gromov [23] in 1999 the idea of group approximation has gained more and more interest, centering around sofic and hyperlinear groups. Especially the work of Elek and Szabó in [14], [15], [16] and [17] helped a lot in understanding the significance and behaviour of sofic groups. A group is sofic if there exist enough *almost homomorphisms* into symmetric groups S_n, in which case we also say that the group can be *approximated* by permutations. Such an almost homomorphism is a mapping which is close to being a homomorphism on a finite set, where closeness is measured using the Hamming metric on S_n. Since properties of sofic groups are in a very weak sense reflected by properties of permutation groups, the study of sofic groups admits the use of tools which can be applied to permutation groups. In particular, because every permutation acts as a permutation matrix on a finite-dimensional vector space, methods of linear algebra and functional analysis help in the study of sofic groups. Conversely, areas of mathematics providing a means for the investigation of sofic groups benefit from these groups. The reason for this is on the one hand that sofic groups are flexible enough to allow for a treatment resulting in the realization of several conjectures holding for sofic groups. On the other hand to date there are no groups known to be non-sofic. Thus from a practical point of view these conjectures become theorems for all groups of reasonable interest. We want to mention here Gottschalk's Surjunctivity Conjecture proved by Gromov [23] and Kaplansky's Direct Finiteness Conjecture proved by Elek and Szabó [14].

Here this thesis has its starting point: Since every symmetric group on a finite set embeds into the general linear group of a finite-dimensional vector space over any field, we consider approximation of groups with matrices instead of permutations. The key idea is to use the function $\ell_{\mathrm{r}}(A) = \frac{\mathrm{rk}(1-A)}{n}$ to measure the *length* of an $n \times n$-matrix A. (We shall use *length functions* on groups instead of metrics, similarly to using e.g. norms on linear spaces.) With this notion of length, approximation with matrices can be done analogously to approximation with permutations in the case of sofic group. We call groups which can be approximated in this way *linearly sofic*, or K-*sofic*, when we consider matrices over a fixed field K. Linearly sofic groups are of twofold interest. On the one hand they admit a treatment using linear algebra, similarly to sofic groups (although in positive characteristic the methods of functional analysis are of little value.) On the other hand, as a natural generalization of sofic groups, they could possibly be a first step towards an example of a non-sofic group. In this direction previously so called *hyperlinear groups* and *weakly sofic* groups were considered. The former are groups which can be approximated using unitary complex matrices and the Hilbert-Schmidt metric (see [37]) and the latter are group which can be approximated with finite groups endowed with any bi-invariant metric (see [22].) Approximation of groups using matrices is also the subject of [3]. As a preparation for the treatment of linearly sofic groups, in Section 2 we consider the above length function ℓ_{r} and a projective version ℓ_{J} defined modulo the center of $\mathrm{GL}_n(K)$. We introduce the *conjugacy length function* ℓ_{c} defined on any finite group, measuring the logarithmic size of conjugacy classes, and show that most of these length functions (and the Hamming length function) compare asymptotically with Lipschitz-type estimates in groups of growing size.

As it turns out, approximation of groups can be seen from two different standpoints, using almost homomorphisms and metric ultraproducts, respectively. While these are practically equivalent, the former has the advantage of being a convenient approach good to work with, whereas the latter is interesting from a model theoretical point of view. The connection between both methods has its origin in a theorem of Elek and Szabó [15], stating that a group is sofic if and only if it embeds into a metric ultraproduct of symmetric groups. A metric ultraproduct of groups is the quotient of an ultraproduct by the maximal normal subgroup of *infinitesimal* elements, i.e. elements of zero distance to the unit element. (Here the length in the ultraproduct is measured by a function obtained as the pointwise ultralimit of length functions.) Hence with group approximation in mind, these infinitesimal elements are of no interest. Nevertheless, as interesting objects on their own, ultraproducts of metric groups offer the chance to be studied by investigating their normal subgroup of infinitesimal elements as well as the metric ultraproduct resulting as a quotient thereof. While Part I deals with linearly sofic groups, hence metric ultraproducts, Part II concentrates on the lattice of normal subgroups of ultraproducts.

When approximating groups with matrices we encounter problems not existent in the study of sofic groups. The most obvious is that matrices depend on a specific field, while permutations, seen as permutation matrices, can be considered over any field. As one of the more convenient fields, here the complex numbers admit a diversity of methods, ranging from elementary topology to the theory of von Neumann algebras. If our field of definition has positive characteristic, this way is blocked. In Section 4 we want to motivate, using the methods of [12], how continuous von Neumann regular rings and continuous geometries can offer a substitute. Although this approach seems not very fashionable nowadays and lacks topological benefits present in functional analysis, we manage to prove that certain metric ultraproducts of matrix groups are not isomorphic. Apart from this special application, we develop basic properties of linearly sofic groups in Section 3. These include the observation that groups which can be approximated using the length function ℓ_{r} or ℓ_{J} can be approximated using the other one. We also show that if K and L are fields of the same characteristic, then being K-sofic or L-sofic are equivalent conditions. The most important result in Section 3 is the insight that the property of being K-sofic is preserved under the group theoretic constructions of direct and inverse limits, extensions by amenable groups, and direct and free products. These permanence properties partly rely on the ability to force almost homomorphisms to take values of a certain length, a method called *amplification*. The analogous conclusions were already known for sofic groups (see [16],) with the even stronger implication that free products of sofic groups amalgamated over amenable groups are sofic (see [11], [17] and [35].) The results in Section 3 are in some parts inspired by or parallel to those found by Arzhantseva and Păunescu in [3], offering a systematic approach to $\mathbb{C}$-sofic groups. Of particular value is the recent description of Jordan normal forms of tensored matrices by Iima and Iwamatsu [27], also used in [3].

The connection of Part I and Part II lies in the fact that projective special linear groups over finite fields constitute a considerable part of finite simple groups. We are one the one hand interested in the structure of the maximal normal subgroup in ultraproducts of these groups and on the other hand know from [1] and [18] that the lattice of normal subgroups of ultraproducts of alternating groups A_n is linearly ordered by inclusion. Based on the classification of finite simple groups, in Section 5 we investigate the lattice of normal subgroups of ultraproducts of arbitrary finite simple groups, to conclude that it is linearly ordered. The essential ingredient in the proof is the main theorem in [31] by Liebeck and Shalev, which allows us to compare the abovementioned length functions ℓ_J and ℓ_c on all finite simple groups deriving from matrix groups.

A corollary of the Peter-Weyl theorem is the observation that all compact simple groups are either finite or compact connected Lie groups. In Section 6 we generalize the results of Section 5 to ultraproducts of compact simple groups by investigating the remaining case of Lie groups. The approach here is again a comparison of two ways of measuring lengths, one of which is essentially by means of the l^1-norm on unitary complex matrices, while the other one is by averaging over angles in maximal tori of Lie groups. The latter method was developed by Nikolov and Segal in [34]. In contrast to the case of finite groups, the situation for Lie groups of unbounded rank is more complicated, due to their nature as continuous objects. We have to add surprisingly intricate arguments from graph theory (see [2] and [20]) and analysis (see [19]) to eventually deduce that the lattice of normal subgroups of compact connected simple Lie groups is distributive; it is linearly ordered if and only if there is a bound on the rank of all Lie groups involved.

Sections 2, 5 and 6 are in large parts identical with [40], although some material is presented in more detail.

Part I

Linear approximation of groups

2 LENGTH FUNCTIONS

§ 1 LENGTH FUNCTIONS AND METRICS

In a group G we write g^G for the conjugacy class of an element g. The group generated by g is $\langle g\rangle$, the group generated by a subset $S \subset G$ is denoted by $\langle S\rangle$, and consequently the normal subgroup generated by g (the normal closure) is $\langle g^G\rangle$. When the group in which conjugation takes place is known, we write $C(g)$ for the conjugacy class of g and $N(g)$ for the normal closure of g. The centralizer of g in G is $C_G(g)$. For a natural number n, the elementwise power of $S \subset G$ is denoted by $S^{\bullet n} := \{g^n \mid g \in S\}$.

Let G be a group. A function

$$\ell : G \to [0, \infty[$$

is called a **length function** on G if for all $g, h \in G$

LF1 $\ell(g) = 0$ if and only if $g = 1$,

LF2 $\ell(g) = \ell(g^{-1})$,

LF3 $\ell(gh) \leq \ell(g) + \ell(h)$.

If moreover $\ell(hgh^{-1}) = \ell(g)$ holds, then we call ℓ **invariant**. If we replace LF1 by $\ell(1) = 0$, then ℓ is a **pseudo length function**.

In fact the notion of pseudo length functions is just a reformulation of the notion of group pseudo metrics. A pseudo metric d on a group G is called **left-invariant** if $d(fg, fh) = d(g, h)$ holds for all g, h and f in G. Analogously d is **right-invariant** if $d(gf, hf) = d(g, h)$ is true, and **bi-invariant** if it is both left-invariant and right-invariant.

Compare this for example to the situation in normed linear spaces:

PROPOSITION 2.1 *Let ℓ be a (pseudo) length function on a group G. Then $d(g, h) := \ell(gh^{-1})$ defines a (pseudo) metric on G. If ℓ is invariant then d is bi-invariant. If conversely d is a (pseudo) metric on G then $\ell(g) := d(g, 1)$ is a (pseudo) length function. If d is bi-invariant then ℓ is invariant.*

Proof. Given ℓ it is clear from the definition that d is a (pseudo) metric. If $\ell(hgh^{-1}) = \ell(g)$ holds, then $d(fg, fh) = \ell(fgh^{-1}f^{-1}) = \ell(gh^{-1}) = d(g,h)$ and right-invariance follows similarly. The reasoning for the second implication is the same. □

We shall tacitly assume that all pseudo length functions we are dealing with are invariant, unless the contrary is explicitly stated.

We say that a group G has **diameter** D with respect to ℓ if $\sup_{g\in G} \ell(g) = D$. This notion coincides with the diameter of metric spaces. We will usually normalize the length functions on groups with bounded diameter to take values in the interval $[0,1]$, i.e. assume that the group in question has diameter 1.

We are interested in the interaction of pseudo length functions and quotient groups.

PROPOSITION 2.2 *Let H be a normal subgroup of the group G with invariant pseudo length function ℓ. Then*

$$\ell_{G/H}(gH) := \inf_{x\in H} \ell(gx)$$

defines an invariant pseudo length function on the group G/H. If G is finite and ℓ is a length function, then $\ell_{G/H}$ is a length function.

Proof. Since ℓ is invariant and H a normal subgroup

$$\ell_{G/H}(gh) = \inf_{x\in H} \ell(ghx) = \inf_{x\in H} \ell(hgx) = \ell_{G/H}(hg).$$

The triangle inequality follows from

$$\begin{aligned}
\ell_{G/H}(gH \cdot hH) &= \inf_{x,y\in H} \ell(ghxy) \\
&= \inf_{x,y\in H} \ell(gx \cdot hy) \\
&\leq \inf_{x,y\in H} \ell(gx) + \ell(hy) \\
&= \ell_{G/H}(gH) + \ell_{G/H}(hH).
\end{aligned}$$

Symmetry and $\ell_{G/H}(1) = 0$ are clear by the definition. Under the additional assumptions the values $\ell_{G/H}$ takes on $(G/H) \setminus \{1\}$ are positive as infima over finitely many positive numbers. □

Note that the previous proposition holds in a more general form for compact groups, as explained in [36], Lemma 4.5.2.

PROPOSITION 2.3 *Let G be a group with invariant pseudo length function ℓ. Then the set*

$$N := \{g \in G \mid \ell(g) = 0\}$$

is a normal subgroup of G. The function $\ell_{G/N}$ is an invariant length function on G/N and N is the smallest normal subgroup with this property.

Proof. It is clear that N is a subgroup of G. By the invariance of ℓ it is also normal. The very definition of N implies that $\ell_{G/N}$ is a length function. Assuming the existence of another normal subgroup M such that $\ell_{G/M}$ is a length function and $g \in N \setminus M$, we derive

$$\ell_{G/M}(gM) = \inf_{x \in M} \ell(gx) \leq \ell(g) = 0,$$

a contradiction. Thus $N \subset M$ and the minimality of N is proved. □

The proof of the following statement is obvious.

PROPOSITION 2.4 *Let G be a group with normal subgroup H and ℓ a pseudo length function on G/H. Then*

$$\ell^G(g) := \ell(gH)$$

defines a pseudo length function on G. If ℓ is invariant, then ℓ^G is invariant, too.

It will turn out to be necessary to study the asymptotics of pseudo length functions on given groups of, in a certain sense, growing size. Let $\mathcal{G} = \{G_n \mid n \in \mathbb{N}\}$ be a countably infinite family of groups with generic length functions ℓ_1 and ℓ_2 defined for every $G \in \mathcal{G}$. We call ℓ_1 **asymptotically bounded** by ℓ_2 if there are constants c and N such that for every $n \geq N$ and every choice of elements $g \in G_n$

$$\ell_1(g) \leq c\ell_2(g).$$

The constant c is called a **modulus** of asymptotic boundedness. The function ℓ_1 is **locally asymptotically bounded** by ℓ_2 in **radius** δ, if the same holds for all $g \in G_n$ satisfying $\ell_1(g) < \delta$, for some $\delta > 0$ not depending on n. We call ℓ_1 and ℓ_2 (**locally**) **asymptotically equivalent** if ℓ_i is (locally) asymptotically bounded by ℓ_j for any choice of $i, j \in \{1, 2\}$.

§ 2 THE CONJUGACY LENGTH

An example of a pseudo length function that can be defined on any finite group G is the **conjugacy length**

$$\ell_c(g) := \frac{\log|C(g)|}{\log|G|}.$$

Since this definition is void of any geometrical interpretation, it is usually hard to compute the conjugacy length. This obstacle will be overcome later when we show that in the groups of primary interest the conjugacy length is closely related to other, easier accessible length functions. Note that the base of the logarithm is not important in this definition.

PROPOSITION 2.5 *Let G be a finite group. Then the function $\ell_c : G \to [0,1]$ is an invariant pseudo length function on G.*

P r o o f. Obviously ℓ_c only takes values in the interval $[0,1]$. Since the conjugacy class of 1 has one element, $\ell_c(1) = 0$. Conjugacy classes of mutually inverse elements have the same size, which implies $\ell_c(g) = \ell_c(g^{-1})$. The conjugacy class of gh is contained in the product of conjugacy classes $C(g)C(h)$. By the functional equation of the logarithm the triangle inequality follows, whence ℓ_c is a pseudo length function. By definition it is invariant under conjugation. □

Note that ℓ_c is a length function if and only if G is centerless, in particular if G is non-abelian and simple. More explicit is the following statement.

PROPOSITION 2.6 *Let G be a finite group. Then*

$$\ell_c(g) = (\ell_c)_{G/Z(G)}(gZ(G))$$

holds for all $g \in G$.

P r o o f. It is not hard to observe that $|C(gz)| = |C(g)|$ for any central element z, which proves

$$(\ell_c)_{G/Z(G)}(gZ(G)) = \inf_{z \in Z(G)} \ell_c(gz) = \inf_{z \in Z(G)} \ell_c(g) = \ell_c(g).$$ □

The next statement is of independent interest and will not be used in what follows.

LEMMA 2.7 *Let G be a finite group. Then for all $g \in G$ and $n \in \mathbb{N}$ the estimate*

$$\ell_c(g^n) \leq \ell_c(g)$$

holds.

Proof. Let h be an arbitrary element in the conjugacy class of g^n, say $h = xg^n x^{-1}$. Then $h = (xgx^{-1})^n \in C(g)^{\bullet n}$. Because $|C(g)^{\bullet n}| \leq |C(g)|$, the claim follows. □

We will proceed by introducing several classes of groups and associated length functions.

§ 3 LENGTH FUNCTIONS ON PERMUTATION GROUPS

We denote by $[n]$ the set consisting of the natural numbers $1, \ldots, n$. The **Hamming length** of a permutation $\pi \in S_n$ is defined as

$$\ell_H(\pi) := \frac{n - |\{i \in [n] \mid \pi(i) = i\}|}{n}.$$

PROPOSITION 2.8 *The function ℓ_H is an invariant pseudo length function on the symmetric group S_n.*

Proof. It is clear that $\ell_H(1) = 1$ and $\ell_H(\pi) = \ell_H(\pi^{-1})$ for all $\pi \in S_n$. The invariance of ℓ_H follows from

$$\begin{aligned}
\ell_H(\sigma^{-1}\pi\sigma) &= \frac{n - \left|\left\{i \in [n] \,\middle|\, \sigma^{-1}\pi\sigma(i) = i\right\}\right|}{n} \\
&= \frac{n - |\{i \in [n] \mid \pi\sigma(i) = \sigma(i)\}|}{n} \\
&= \frac{n - |\{i \in [n] \mid \pi(i) = i\}|}{n} = \ell_H(\pi).
\end{aligned}$$

The triangle inequality is obtained by estimating

$$\begin{aligned}\ell_{\mathrm{H}}(\pi\sigma) &= \frac{n - \left|\{i \in [n] \,\middle|\, \pi(i) = \sigma^{-1}(i)\}\right|}{n} \\ &\leq \frac{2n - |\{i \in [n] \,|\, \pi(i) = i\}| - |\{i \in [n] \,|\, \sigma(i) = i\}|}{n} \\ &= \ell_{\mathrm{H}}(\pi) + \ell_{\mathrm{H}}(\sigma).\end{aligned}$$

Note that the same construction works when we replace $[n]$ equipped with the counting measure by an arbitrary propability measure space and S_n by the group of all measure preserving transformations. (Confer [37], Example 2.2.)

Invariant length functions necessarily have to be constant on conjugacy classes. In this sense the following example qualifies as a putative invariant length function on symmetric groups.

PROPOSITION 2.9 *Let π be a permutation in S_n with l cycles. Then*

$$\ell_{\mathrm{r}}(\pi) := \frac{n - l}{n}$$

defines an invariant length function on S_n.

We postpone the proof to § 5.

The following proposition serves as an introductory example of asymptotic equivalence and will be useful later.

PROPOSITION 2.10 *The length functions ℓ_{H} and ℓ_{r} are asymptotically equivalent in $\mathcal{S}$.*

Proof. Let $\pi \in S_n$ with l cycles, m of which are trivial. Then immediately

$$\ell_{\mathrm{r}}(\pi) = \frac{n - l}{n} \leq \frac{n - m}{n} = \ell_{\mathrm{H}}(\pi)$$

follows. Because the remaining $l - m$ non-trivial cycles have length at least 2, $l - m \leq \frac{1}{2}(n - m)$. We conclude

$$\frac{n - m}{n} = \frac{n - l + l - m}{n} \leq \frac{n - l}{n} + \frac{n - m}{2n}$$

and finally $\ell_{\mathrm{H}}(\pi) \leq 2\ell_{\mathrm{r}}(\pi)$. □

We shall use the remainder of this paragraph to show the asymptotic equivalence of the Hamming length and the conjugacy length introduced above.

LEMMA 2.11 *Let π be a permutation in S_n. If the number of cycles of length i is denoted by c_i and the longest cycle has length k, then the cardinality of the conjugacy class of π in S_n is given by*

$$|C(\pi)| = n!\left(\prod_{i=1}^{k} i^{c_i} \prod_{i=1}^{k} c_i!\right)^{-1}.$$

Proof. The claim is elementary and follows by combinatorics as explained in [44], Section 2.3.1. □

LEMMA 2.12 *In $\mathcal{S}$ the length function ℓ_c is asymptotically bounded by ℓ_H.*

Proof. We consider a non-trivial permutation $\pi \in S_n$ with m fixed points. Again we denote the number of cycles of length i of π by c_i. Then by assumption $c_1 = m$. By Stirling's formula for large n the estimate

$$\tfrac{1}{2} n \log n \le \log n! \le 2n \log n$$

holds.

Using Lemma 2.11 we obtain the trivial inequality

$$|C(\pi)| \le \frac{n!}{m!}.$$

Thus

$$\frac{\sum_{i=1}^{n} \log i - \sum_{i=1}^{m} \log i}{\frac{1}{2} n \log n} \le 2\frac{\sum_{i=m+1}^{n} \log n}{n \log n} = 2\frac{n-m}{n}$$

implies

$$\ell_c(\pi) \le 2\ell_H(\pi),$$

which concludes the proof. □

The short proof of the following lemma we owe to Nikolay Nikolov.

LEMMA 2.13 *In $\mathcal{S}$ the length function ℓ_H is asymptotically bounded by ℓ_c.*

P r o o f. Let π be a permutation in S_n. Assume that π has $n-k$ fixed points, i.e. $\ell_{\mathrm{H}}(\pi)=\frac{k}{n}$. For the sake of simplicity we only treat the case of even k and note that the odd case is almost the same. We distinguish the cases $k>\frac{1}{2}n$ and $k\leq\frac{1}{2}n$.

If $k>\frac{1}{2}n$ we can estimate the size of the centralizer of π by

$$|C_{S_n}(\pi)|\leq n(n-1)\cdot\ldots\cdot(n-\tfrac{1}{2}k),$$

since π has at most $n-\frac{1}{2}k$ cycles and every permutation commuting with π is determined by its action on one point from each cycle of π. Therefore

$$|C(g)|\geq\frac{n!}{n(n-1)\cdot\ldots\cdot(n-\frac{1}{2}k)}=(\tfrac{1}{2}k)!.$$

By loosely applying Stirling's approximation $(\frac{1}{2}k)!\geq(\frac{1}{2}k)^{\frac{1}{4}k}$ follows. Since $\log(\frac{1}{2}k)\geq\frac{1}{2}\log n$ for $n\geq 17$

$$\frac{\log|C(\pi)|}{\log(n!)}\geq\frac{\log((\frac{1}{2}k)!)}{\log(n!)}\geq\frac{\frac{1}{2}\log(\frac{1}{2}k)\frac{k}{2}}{n\log n}\geq\frac{k}{8n}.$$

This means $\ell_{\mathrm{H}}(\pi)\leq 8\ell_{\mathrm{c}}(\pi)$.

If $k\leq\frac{1}{2}n$, then π has at most $\frac{1}{2}k$ non-trivial cycles. Since the permutations commuting with π are determined by the action of a single point from each cycle, we deduce the estimate

$$|C_{S_n}(\pi)|\leq(n-k)!\cdot k^{\frac{1}{2}k}.$$

It is clear that $k^{\frac{1}{2}k}\leq(\frac{1}{2}n)^{\frac{1}{2}k}$ and $n(n-1)\cdot\ldots\cdot(n-k+1)\geq(\frac{1}{2}n)^k$, and therefore

$$\begin{aligned}|C(\pi)|&\geq\frac{n!}{(n-k)!\cdot k^{\frac{1}{2}k}}\\&\geq\frac{n(n-1)\cdot\ldots\cdot(n-k+1)}{k^{\frac{1}{2}k}}\\&\geq\frac{(\frac{1}{2}n)^k}{(\frac{1}{2}n)^{\frac{1}{2}k}}=(\tfrac{1}{2}n)^{\frac{1}{2}k}.\end{aligned}$$

Because $2\log(\frac{1}{2}n)\geq\log n$, we finally obtain

$$\ell_{\mathrm{c}}(g)\geq\frac{\frac{k}{2}\log(\frac{1}{2}n)}{\log(n!)}\geq\frac{k\log(\frac{1}{2}n)}{2n\log n}\geq\tfrac{1}{4}\ell_{\mathrm{H}}(\pi).$$

□

THEOREM 2.14 *In $\mathcal{S}$ and $\mathcal{A}$ alike the length functions ℓ_{H} and ℓ_{c} are asymptotically equivalent.*

P r o o f. The claim concerning $\mathcal{S}$ follows from Lemma 2.12 and Lemma 2.13. The conjugacy classes of S_n behave in two different ways. Either they correspond to exactly one conjugacy class in A_n, or they split into two classes in A_n. In the first case the size of the conjugacy class stays the same, whereas in the second case it splits into two parts of equal size. (Confer [44], Paragraph 2.3.2.) Now again Lemma 2.12 and Lemma 2.13 apply. □

§ 4 RANK, NORM AND LENGTH

Many interesting examples of groups arise as the groups of invertible elements of rings and algebras or subgroups thereof. Moreover of special interest for us are length functions that can be defined using additional structure of the underlying ring or algebra.

Let A be a unitary ring. We call a function $R\colon A \to [0,1]$ a **rank function** or simply **rank** if

RF1 $R(a) = 0$ if and only if $a = 0$,

RF2 $R(1) = 1$,

RF3 $R(a) = R(-a)$,

RF4 $R(a+b) \leq R(a) + R(b)$,

RF5 $R(ab) \leq \min(R(a), R(b))$.

A function $N\colon A \to [0,\infty[$ is a **norm** if RF5 in the definition of rank functions is replaced by

NF5 $N(ab) \leq N(a)N(b)$.

We define **pseudo rank functions** and **pseudo norms** by allowing for $R(a) = 0$ and $N(a) = 0$ even if $a \neq 0$. In what follows all rings are assumed to be unitary. The proof of the following propositions is a straightforward application of the definitions.

PROPOSITION 2.15 *Let A be a ring with (pseudo) rank or norm P. Then*

$$d(a,b) := d_P(a,b) := P(a-b)$$

defines a (pseudo) metric on A.

PROPOSITION 2.16 *Let A be a ring with pseudo rank or norm P. Then the set*

$$N_P := \{a \in A \,|\, P(a) = 0\}$$

is an ideal in A.

PROPOSITION 2.17 *Let A be a ring with pseudo rank or norm P. Then*

$$\overline{P}(a + N_P) := P(a)$$

defines a rank or norm, respectively, on the quotient ring A/N_P.

In any ring A we write $A^\times$ to denote the group of invertible elements in A. If A is a ring with a (pseudo) rank or norm P, then we denote the set of invertible elements g in A satisfying $P(g) = P(g^{-1}) = 1$ by $A_1^\times$. The set $A_1^\times$ is called the **rank group** or **norm group**, respectively, and the justification for this notion follows immediately.

PROPOSITION 2.18 *Let A be a ring with (pseudo) rank function R. Then*

$$A_1^\times = \left\{g \in A_1^\times \,\middle|\, R(g) = 1\right\} = A^\times$$

and in particular $A_1^\times$ is a group.

Proof. The inclusions from left to right are clear. Let $g \in A^\times$ and observe

$$1 = R(1) = R(g\,g^{-1}) \leq \min(R(g^{-1}), R(g)).$$

Then by definition $1 \leq R(g), R(g^{-1}) \leq 1$ and we have proved $A^\times \subset A_1^\times$. $\square$

PROPOSITION 2.19 *Let A be a ring with (pseudo) norm N. Then $A_1^\times \subset A^\times$ is a group.*

P r o o f. If $g, h \in A_1^\times$, then

$$\begin{aligned} N(g) &= N(ghh^{-1}) \le N(gh)N(h^{-1}) \\ &= N(gh) \le N(g)N(h). \end{aligned}$$

Since the left and right side both equal 1, we deduce $N(gh) = 1$ and thus $gh \in A_1^\times$. □

PROPOSITION 2.20 *Let A be a ring with (pseudo) rank or norm P. Then*

$$\ell(g) := \ell_P(g) := P(1-g)$$

defines an invariant (pseudo) length function on $A_1^\times$.

P r o o f. Let $g, h \in A_1^\times$ be arbitrary. It is clear that $\ell(1) = 0$ holds. If $P(g) = 0$ implies $g = 0$ and $\ell(g)$ equals 0, then $P(1-g) = 0$ implies $1 - g = 0$ and consequently $g = 1$. Assume that P is a norm. Then we see

$$\begin{aligned} \ell(g) &= P(1-g) = P(g(g^{-1}-1)) \\ &\le P(g)P(g^{-1}-1) \le P(1-g^{-1}) = \ell(g^{-1}) \end{aligned}$$

and by symmetry $\ell(g^{-1}) \le \ell(g)$, and so $\ell(g) = \ell(g^{-1})$ follows. Further

$$\begin{aligned} \ell(gh) &= P(1-gh) \le P(h^{-1}-g)P(h) \\ &\le P(h^{-1})P(1-hg) \le P(1-hg) = \ell(hg) \end{aligned}$$

and by symmetry $\ell(gh) = \ell(hg)$. Hence follows invariance of ℓ. At last

$$\begin{aligned} \ell(gh) &= P(1-gh) \\ &= P(1-h+h-gh) \\ &\le P(1-h) + P((1-g)h) \\ &= P(1-h) + P(1-g) \\ &= \ell(g) + \ell(h) \end{aligned}$$

proves the triangle inequality.

The computations work equally well for rank functions. □

Note that the (pseudo) metric group thus obtained has diameter not more than 2 and we shall normalize the length to achieve diameter 1 in most applications. It is noteworthy that for norms the restriction to elements with norm 1 is necessary. Consider for example the ring of complex 2×2-matrices with the operator norm. The matrix

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$$

is invertible and not of norm 1, and $\ell(A) = \frac{1}{2}||1-A|| = 1$ and $\ell(A^2) = 4$. Hence the triangle inequality would yield $4 = \ell(A^2) \leq 2\ell(A) = 2$.

PROPOSITION 2.21 *Let A be a ring with (pseudo) rank or norm P. Then the topology on $A_1^\times$ induced by ℓ_P coincides with the topology inherited from A.*

Proof. This follows immediately from the identity

$$P(a-b) = P(1-ba^{-1}) = \ell_P(ba^{-1}),$$

holding for all $a, b \in A_1^\times$. □

We conclude the development of length functions obtained from rings with rank or norm with the following theorem.

THEOREM 2.22 *Let A be a ring with pseudo rank or norm P. We denote the subgroup of all elements g in $A_1^\times$ satisfying $\ell_P(g) = 0$ by $N := N_{\ell_P}$. Then*

$$(A/N_P)_1^\times \cong A_1^\times / N_{\ell_P}$$

and

$$\ell_{\overline{P}} = (\ell_P)_{A_1^\times/N}.$$

Proof. We define

$$\varphi\colon (A/N_P)_1^\times \to A_1^\times/N_{\ell_P}, \qquad g + N_P \mapsto gN_{\ell_P}.$$

Because $g + N_P = h + N_P$ implies

$$\ell_P(gh^{-1}) = P(1-gh^{-1}) = P(g-h) = 0$$

and hence $gN_{\ell_P} = hN_{\ell_P}$, φ is well defined. The same line of thought shows that φ is injective and it is clear that φ is surjective. The calculations to show that φ is a homomorphism are standard. At last we compute

$$\begin{aligned}\ell_{\overline{P}}(g+N_P) &= \overline{P}(1-g+N_P) = P(1-g)\\ &= \ell_P(g) = (\ell_P)_{A_1^\times/N}(gN)\end{aligned}$$

to conclude the proof. □

As suggested by the notions introduced so far, the examples of primary interest are rings of matrices and rings of operators. For an instance consider the ring $M_n(\mathbb{C})$ of complex $n \times n$-matrices. On this ring we have the usual norm $||a||_1 = \sum_{i,j=1}^n |a_{ij}|$. The norm group coincides with the group $\mathrm{U}_n(\mathbb{C})$ of unitary matrices. The above reasoning allows us to conclude that

$$\ell_1(g) := \frac{||1-g||_1}{2n}$$

is an invariant length function on $\mathrm{U}_n(\mathbb{C})$.

In this vein we present another length function

$$\ell_2(g) := \frac{||1-g||_2}{2n},$$

where $||a|| = \left(\sum_{i,j=1}^n |a_{ij}|^2\right)^{\frac{1}{2}}$ is the Hilbert-Schmidt norm.

Investigation of the typical example of pseudo length functions obtained from a rank function is the purpose of the next paragraph.

§ 5 LENGTH FUNCTIONS ON LINEAR GROUPS

Given a finite-dimensional vector space V we write $\mathrm{GL}(V)$ for all bijective linear transformations of V and $\mathrm{SL}(V)$ for all linear transformations of V of determinant 1. When $V = K^n$ for some field K we use notation $\mathrm{GL}_n(K) := \mathrm{GL}(V)$ and the like, which reduces further to $\mathrm{GL}_n(q) := \mathrm{GL}_n(\mathbb{F}_q)$ etc., when K is the finite field $\mathbb{F}_q$ of cardinality q. We will think of elements in $\mathrm{GL}_n(K)$ as matrices corresponding to the standard basis in K^n.

We shall deal in particular with linear groups over finite fields and introduce the symbols $\mathcal{GL}(q)$ for the class of all general linear groups defined over the field $\mathbb{F}_q$ and $\mathcal{GL}_{\text{fin}}$ for the union of these, where q ranges over all prime powers. Exchanging general linear groups with special linear groups yields $\mathcal{SL}(q)$ and $\mathcal{SL}_{\text{fin}}$. If V is a vector space over a field K we will write 1 for the identical mapping $V \to V$ and write simply α for the mapping $\alpha \cdot 1$, where $\alpha \in K$.

PROPOSITION 2.23 *Let V be a vector space of dimension n. Then*

$$\ell_{\mathrm{r}}(g) := \frac{\mathrm{rk}(1-g)}{n}$$

is an invariant length function on $\mathrm{GL}(V)$.

P r o o f. This is a corollary of Proposition 2.20 □

We call the function ℓ_{r} the **rank length**.

Now the following conclusion is rather obvious.

P r o o f of Proposition 2.9. The symmetric group S_n embeds as the subgroup of permutation matrices into $\mathrm{GL}_n(K)$, where K is any field. If π consists of the cycles $\pi_1, \dots, \pi_l$, then the corresponding permutation matrix P_π equals the direct sum $P_{\pi_1} \oplus \dots \oplus P_{\pi_l}$. Therefore $\mathrm{rk}(\mathrm{id} - \pi_i) = k - 1$ if π_i has length k. Hence ℓ_{r} is the restriction of the rank length to permutations. □

We want to replace the conjugacy length in general linear groups over finite fields with a length function permitting more geometric insight, comparable to the situation of permutation groups in § 3. As it turns out it is necessary to gain some independence of the base field. We therefore introduce the **Jordan length** (the name of which is explained below).

$$\ell_{\mathrm{J}}(g) = \inf_{\alpha \in K^\times} \frac{\mathrm{rk}(\alpha - g)}{n}.$$

As the center of $\mathrm{GL}(V)$ is isomorphic to $K^\times$, by Proposition 2.2 and Proposition 2.4 it is clear that ℓ_{J} is a pseudo length function. From now on we shall write

$$m_g := \sup_{\alpha \in K^\times} \dim(\ker(\alpha - g)),$$

whenever g is an element in a linear group over a field K. With this definition another characterization of the Jordan length is

$$\ell_{\mathrm{J}}(g) = \frac{n - m_g}{n}.$$

PROPOSITION 2.24 *Let V be a finite-dimensional vector space and g an element in* $\mathrm{GL}(V)$. *If $\ell_{\mathrm{r}}(g) \leq \delta$, then $\ell_{\mathrm{J}}(g) \geq \min\{(1-\delta), \delta\}$.*

Proof. Let $m = \mathrm{rk}(1-g)$. In the easiest case $\ell_{\mathrm{J}}(g) = \ell_{\mathrm{r}}(g) \geq \delta$. Hence we can assume $m \neq m_g$. Then of course $m + m_g \leq n$ and

$$\begin{aligned} \ell_{\mathrm{J}}(g) &= \frac{n - m_g}{n} \geq \frac{n-(n-m)}{n} \\ &= 1 - \ell_{\mathrm{r}}(g) \geq 1 - \delta \end{aligned}$$

follows. □

COROLLARY 2.25 *Let g be an element in* $\mathrm{GL}(V)$. *If $\ell_{\mathrm{r}}(g) \leq \frac{1}{2}$, then $\ell_{\mathrm{r}}(g) = \ell_{\mathrm{J}}(g)$.*

Proof. By definition $\ell_{\mathrm{J}}(g) \leq \ell_{\mathrm{r}}(g)$. □

We call an element $A \in \mathrm{GL}_n(K)$ **semisimple** if A is diagonalizable in $\mathrm{GL}_n(\overline{K})$, where $\overline{K}$ is the algebraic closure of K. A matrix A is **unipotent** if all its eigenvalues are equal to 1.

We cite from the introduction of the Jordan decomposition on pp. 395, 396 in [31]. Every $A \in \mathrm{GL}_n(q)$ equals the commuting product $A_{\mathrm{s}}A_{\mathrm{u}}$ of a semisimple matrix A_{s} and a unipotent matrix A_{u}, this being called the **Jordan decomposition**.

We denote by J_k the unipotent $k \times k$-Jordan matrix

$$\begin{pmatrix} 1 & 1 & & & \\ & 1 & 1 & & \\ & & \ddots & \ddots & \\ & & & 1 & 1 \\ & & & & 1 \end{pmatrix}.$$

If m_{ij} are non-negative integers for all $j = 1 \dots r$, $i = 1 \dots l_j$, we let $J^{(m_j)} := m_{1j}J_1 \oplus m_{2j}J_2 \oplus \dots \oplus m_{l_j j}J_{l_j}$, where $mJ_i := J_i \oplus \dots \oplus J_i$ is the direct sum of m Jordan blocks of size i. Then A is conjugate to a matrix

$$\alpha_1 J^{(m_1)} \oplus \dots \oplus \alpha_r J^{(m_r)} \oplus \lambda_1 \otimes J^{(n_1)} \oplus \dots \oplus \lambda_t \otimes J^{(n_t)},$$

where m_j and n_j are appropriate finite sequences of non-negative integers, $\alpha_j \in \mathbb{F}_q^\times$, and the λ_j are irreducible matrices. In this representation we can assume that $m := m_{11} \geq m_{12} \geq \dots \geq m_{1r}$. That is m counts the maximal number of Jordan blocks of size 1 to an eigenvalue $\alpha_1 \in \mathbb{F}_q$ of A.

We are interested in the following result.

THEOREM 2.26 *The pseudo length functions ℓ_c and ℓ_J are asymptotically equivalent in $\mathcal{GL}_{\text{fin}}$ and $\mathcal{SL}_{\text{fin}}$.*

Unfortunately the least important part of the theorem takes the greatest effort. Nevertheless we can rely on the weaker Theorem 2.28 when it comes to an application in Section 3. Having proved Theorem 2.28, we shall for the sake of completeness develop the remaining parts of the proof of Theorem 2.26.

LEMMA 2.27 *The orders of $\mathrm{GL}_n(q)$ and $\mathrm{SL}_n(q)$ are more than q^{n^2-n} and less than q^{n^2}.*

Proof. The observation $|\mathrm{SL}_n(q)| \leq |\mathrm{GL}_n(q)| \leq q^{n^2}$ is trivial. The exact order of $\mathrm{SL}_n(q)$ is well known and equals

$$|\mathrm{SL}_n(q)| = q^{\frac{1}{2}n(n-1)}(q^2-1)(q^3-1)\cdot \ldots \cdot (q^n-1).$$

This is more than

$$q^{\frac{1}{2}n(n-1)} \cdot q \cdot q^2 \cdot \ldots \cdot q^{n-1} \geq q^{n(n-1)}. \qquad \square$$

THEOREM 2.28 *In the classes of groups $\mathcal{GL}_{\text{fin}}$ and $\mathcal{SL}_{\text{fin}}$ the pseudo length function ℓ_c is asymptotically bounded by ℓ_J, while ℓ_J is locally asymptotically bounded by ℓ_c in radius $\frac{1}{2}$.*

P r o o f. Let $1 \neq g \in G := \mathrm{GL}_n(q)$ where m_g and $m = m_{11}$ are defined as above. Let α be an eigenvalue such that $\ker(\alpha - g)$ has dimension m_g and β an eigenvalue such that the Jordan decomposition of g shows m Jordan blocks of size 1 corresponding to β. We want to compare m and m_g. By the definition of m_g it is clear that $m \leq m_g$. The number of Jordan blocks of size 1 corresponding to α is not more than m. Since the other Jordan blocks adding to m_g have size at least 2 we conclude $m_g - m \leq \frac{1}{2}(n-m)$. Because $C_G(g)$ contains a subgroup isomorphic to $\mathrm{GL}_m(q)$, $C(g)$ does not have more than $q^{n^2-m^2+m}$ elements, using also Lemma 2.27. We find

$$\begin{aligned} \frac{\log|C(g)|}{\log|G|} &\leq \frac{n^2-m^2+m}{n^2-n} \leq 2\frac{n^2-nm}{n^2} \\ &\leq 4\frac{n-m+m-m_g}{n} = 4\frac{n-m_g}{n}, \end{aligned}$$

and consequently

$$\ell_{\mathrm{c}}(g) \leq 4\ell_{\mathrm{J}}(g).$$

Now assume $\ell_{\mathrm{J}}(g) \leq \frac{1}{2}$, i.e. $m_g \geq \frac{1}{2}n$. From [21], Chapter VIII, Paragraph 2 we know that a matrix in $C_G(g)$ is the direct sum of two blocks of sizes $m_g \times m_g$ and $n - m_g \times n - m_g$, respectively. We calculate

$$\begin{aligned} m_g^2 + (n-m_g)^2 &\leq m_g^2 + \tfrac{1}{2}n(n-m_g) \\ &\leq nm_g + \tfrac{1}{2}n(n-m_g) \\ &= \tfrac{1}{2}nm_g + \tfrac{1}{2}n^2 \end{aligned}$$

to deduce that $C_G(g)$ has at most $q^{\frac{1}{2}n^2+\frac{1}{2}nm_g}$ elements and $C(g)$ more than $q^{\frac{1}{2}n^2-\frac{1}{2}nm_g-n}$. Since $n-m \leq 2(n-m-2)$ for $n \geq 8$,

$$\frac{n-m_g}{n} \leq 2\frac{n^2-nm_g-2n}{n^2} \leq 4\frac{\frac{1}{2}(n^2-nm_g)-n}{n^2} \leq 4\frac{\log|C(g)|}{\log|G|}.$$

This means

$$\ell_{\mathrm{J}}(g) \leq 4\ell_{\mathrm{c}}(g).$$

The same argument is valid when working in special linear groups. □

Proof of Theorem 2.26. One implication and a half of the claim have already been proved.

Lemma 5.4 in [31] states that there is a universal constant c such that whenever $1 \neq g \in \mathrm{SL}_n(q)$ and $k \geq \frac{cn}{n-m}$, then $C(g)^k = \mathrm{SL}_n(q)$. We can assume c an integer and k minimal such that $k = \varepsilon c \frac{n}{n-m_g} \geq \varepsilon c \frac{n}{n-m}$, where the error ε is definitely less than 2. Now $|\mathrm{SL}_n(q)| \leq |C(g)|^k$ and

$$k\ell_{\mathrm{c}}(g) = \frac{\log(|C(g)|^k)}{\log|\mathrm{SL}_n(q)|} \geq \frac{\log|C(g)^k|}{\log|\mathrm{SL}_n(q)|} = 1.$$

This implies

$$\ell_{\mathrm{c}}(g) \geq k^{-1} = \varepsilon^{-1}c^{-1}\frac{n-m_g}{n} \geq \tfrac{1}{2}c^{-1}\ell_{\mathrm{J}}(g).$$

By comparison of the sizes of conjugacy classes in $\mathrm{SL}_n(q)$ and $\mathrm{GL}_n(q)$ the claim follows for $\mathcal{GL}(q)$. Because the above argument is independent of q, we have proved the claim for $\mathcal{SL}_{\mathrm{fin}}$ and $\mathcal{GL}_{\mathrm{fin}}$. □

The last proof implicitly used the hard calculations in [31], designed to work in a broader generality than in $\mathrm{SL}_n(q)$ or $\mathrm{GL}_n(q)$. We dedicate some time to the task of estimating the size of conjugacy classes in $\mathrm{GL}_n(q)$ and thus deriving an elementary proof of Theorem 2.26.

LEMMA 2.29 *Let $s_i \geq 2$ be natural numbers such that $\sum_{i=1}^k s_i = n$. Then $\sum_{i=1}^k (2i-1)s_i \leq \frac{1}{2}n^2$.*

Proof. We first claim that for all possible values of s_1 the estimate $\frac{1}{2}s_1^2 + 2n \leq ns_1 + s_1$ holds. We try induction on s_1, starting with a positive result for $s_1 = 2$. Now we assume $s_1 \geq 3$ and $\frac{1}{2}(s_1-1)^2 + 2n \leq n(s_1-1) + s_1 - 1$. Then

$$\begin{aligned}
\tfrac{1}{2}s_1^2 + 2n &\leq \tfrac{1}{2}s_1^2 + \tfrac{1}{2} - s_1 + n + 2n + 1 \\
&\leq \tfrac{1}{2}(s_1-1)^2 + 2n + n + 1 \\
&\leq n(s_1-1) + s_1 - 1 + n + 1 \\
&= ns_1 + s_1.
\end{aligned}$$

We proceed by induction on k and note that the case $k=1$ works. Assume that $k \geq 2$ and the claim holds for $k-1$. Then

$$\begin{aligned}\sum_{i=1}^{k}(2i-1)s_i &= \sum_{i=1}^{k-1}(2i+1)s_{i+1}+s_1\\ &= \sum_{i=1}^{k-1}(2i-1)s_{i+1}+2\sum_{i=1}^{k-1}s_{i+1}+s_1\\ &\leq \tfrac{1}{2}(n-s_1)^2+2(n-s_1)+s_1\\ &= \tfrac{1}{2}n^2-ns_1+\tfrac{1}{2}s_1^2+2n-s_1 \leq \tfrac{1}{2}n^2,\end{aligned}$$

where the last inequality follows from the first claim in this proof. □

LEMMA 2.30 *The centralizer of a unipotent Jordan matrix J in $\mathrm{GL}_n(q)$ has size at most $q^{\frac{1}{2}n^2+\frac{1}{2}m^2}$, where m is the number of trivial Jordan blocks.*

Proof. For a start assume that J has k Jordan blocks, none of which is trivial (that is $m=0$). Let A be in $\mathrm{GL}_n(q)$ and write $A=(A_{ij})_{1\leq i,j\leq k}$, where block A_{ij} has size $s_i \times s_j$ and s_i denotes the size of the i-th Jordan block of J. If A commutes with J, then for example [21], Chapter VIII, Paragraph 2 (or a direct calculation) shows that the blocks A_{ij} are necessarily of the form

$$\begin{pmatrix} 0 & \dots & 0 & a_1 & a_2 & \dots & \dots & a_{s_i}\\ & & & & a_1 & a_2 & & \vdots\\ \vdots & & & & \ddots & \ddots & \ddots & \vdots\\ & & & & & & a_1 & a_2\\ 0 & & & \dots & & & 0 & a_1 \end{pmatrix}$$

if $s_i \leq s_j$. For $s_j \leq s_i$ the block A_{ij} is in upper triangular form, too. Now in each block there are only $\min(s_i,s_j)$ possible different entries. Thus the size of the centralizer of J is restricted to q^S at most, where $S=\sum_{i,j=1}^{k}\min(s_i,s_j)$. If we assume that $s_1 \geq s_2 \geq \ldots \geq s_k$, then S equals $\sum_{i=1}^{k}(2i-1)s_i$. By Lemma 2.29 this sum amounts to less than $\frac{1}{2}n^2$ and hence $|C_{\mathrm{GL}_n(q)}(J)| \leq q^{\frac{1}{2}n^2}$.

Now we approach the general case and write $J=I_m \oplus J'$, where J' has no trivial Jordan blocks. We assume that A commutes with J and write A as a block matrix $(A_{ij})_{1\leq i,j\leq 2}$, where block A_{11} has size $m\times m$. The equation $JA=AJ$ yields the obstructions $A_{12} = A_{12}J'$, $J'A_{21} = A_{21}$ and $J'A_{22} = A_{22}J'$. The last case is handled by the above conclusion. The second case implies that each column of A_{21} is an eigenvector of J'. Clearly the dimension of the sum of eigenspaces of J' is at most $\frac{1}{2}(n-m)$ and hence the number of possible eigenvectors is bounded above by $q^{\frac{1}{2}(n-m)}$. As A_{21} has m columns, there are only $q^{\frac{1}{2}m(n-m)}$ matrices that can occur as A_{21}. The first case is treated in the same way. All in all the upper bound on the size of the centralizer of J amounts to

$$q^{m^2+2\cdot\frac{1}{2}m(n-m)+\frac{1}{2}(n-m)^2} = q^{\frac{1}{2}n^2+\frac{1}{2}m^2},$$

as claimed. □

LEMMA 2.31 *Let $m_{1j} \leq k_j$, $j = 1\dots r$ be natural numbers such that $k := \sum_{j=1}^{r} k_j \leq n$. Assume moreover that $m_{11} \geq m_{12} \geq \dots \geq m_{1r}$. Then*

$$k_1^2 + \sum_{j=2}^{r}(k_j^2 + m_{1j}^2) \leq 2nk - k^2.$$

P r o o f. The proof is by induction on r. For $r = 1$ the sum equals $k = k_1$ and hence

$$k_1^2 = k^2 = 2k^2 - k^2 \leq 2nk - k^2.$$

Assume that $r \geq 2$ and the claim holds for $r-1$. Then

$$\begin{aligned} k_1^2 + \sum_{j=2}^{r}(k_j^2 + m_{1j}^2) &= k_1^2 + \sum_{j=2}^{r-1}(k_j^2 + m_{1j}^2) + k_r^2 + m_{1r}^2 \\ &\leq 2(n-k_r)(k-k_r) - (k-k_r)^2 + k_r^2 + m_{1r}^2 \\ &= 2nk - k^2 + 2k_r^2 + m_{1r}^2 - 2nk_r. \end{aligned}$$

We show that the last three summands on the right are smaller than 0. By definition $m_{1r} \leq k_r \leq \frac{n}{r}$. Hence

$$\begin{aligned} 2k_r^2 + m_{1r}^2 - 2nk_r &\leq \frac{2n}{r}k_r + \frac{n}{r}k_r - 2nk_r \\ &\leq 2k_r\left(\frac{3n}{r} - 2n\right) \\ &\leq 2k_r n\left(\tfrac{3}{2} - 2\right), \end{aligned}$$

because $r \geq 2$. The claim follows. □

LEMMA 2.32 *In the the class of groups $\mathcal{GL}(q)$, ℓ_{J} is asymptotically bounded by ℓ_{c}.*

Proof. Let g be a non-trivial element in $\mathrm{GL}_n(q)$. Let $g = su$ be the Jordan decomposition of g, where s is semisimple and u unipotent. Then by Section 5 in [31]

$$C_{\mathrm{GL}_n(q)}(g) = C_{\mathrm{GL}_n(q)}(s) \cap C_{\mathrm{GL}_n(q)}(u).$$

Furthermore

$$C_{\mathrm{GL}_n(q)}(s) \cong \prod_{j=1}^{r} \mathrm{GL}_{k_j}(q) \times \prod_{j=1}^{t} \mathrm{GL}_{b_j}(q^{a_j}),$$

where, using the notation as in the introduction of the Jordan decomposition above, every irreducible $a_j \times a_j$-matrix occurs b_j times. The projection of u in $\mathrm{GL}_{k_j}(q)$ is conjugate to a matrix $J^{(m_j)}$, where $\sum_{i=1}^{l_j} i m_{ij} = k_j$. The projection in $\mathrm{GL}_{b_j}(q^{a_j})$ is conjugate to a matrix $J^{(n_j)}$.

Using Lemma 2.30 this decomposition implies that $|C_{\mathrm{GL}_n(q)}(g)| \le q^S$, where

$$S = \tfrac{1}{2}\sum_{j=1}^{r}(k_j^2 + m_{1j}^2) + \sum_{j=1}^{t} a_j b_j^2.$$

The first observation is

$$\sum_{j=1}^{t} a_j b_j^2 \le \sum_{j=1}^{t} a_j b_j \cdot \tfrac{1}{2}(n-k) \le \tfrac{1}{2}(n-k)^2,$$

where we abbreviate $k = \sum_{j=1}^{r} k_j$ as in Lemma 2.31. Therefore by Lemma 2.31

$$\begin{aligned} 2S &\le m_{11}^2 + k_1^2 + \sum_{j=2}^{r}(k_j^2 + m_{1j}^2) + (n-k)^2 \\ &\le m_{11}^2 + 2nk - k^2 + (n-k)^2 \\ &= m_g^2 + n^2. \end{aligned}$$

Thus by Lemma 2.27 the size of the conjugacy class of g is more than $q^{\frac{1}{2}(n^2-n-m_g^2)}$ and consequently

$$\ell_{\mathrm{c}}(g) \ge \frac{n^2 - 2n - m_g^2}{2n^2} \ge \tfrac{1}{4}\ell_{\mathrm{J}}(g)$$

for n large enough. □

Now Theorem 2.28 and Lemma 2.32 suffice to reprove Theorem 2.26.

3 ULTRAPRODUCTS AND APPROXIMATION

§ 1 ULTRAPRODUCTS AND ULTRALIMITS

We give a short introduction to filter theory. Beyond the definitions we shall cite the main theorems without reproduction of the proofs. For an introduction to filter theory confer [6] and for a thorough treatment of ultraproducts and metric ultraproducts [4]. Let X be a set. A subset $\mathfrak{f}$ of the powerset $\mathfrak{P}(X)$ of X is a **filter** if

FL1 $X \in \mathfrak{f}$,

FL2 $\varnothing \notin \mathfrak{f}$,

FL3 $A, B \in \mathfrak{f}$ implies $A \cap B \in \mathfrak{f}$,

FL4 $A \in \mathfrak{f}$ and $A \subset B$ implies $B \in \mathfrak{f}$.

One should think of a filter *in* X as a generalization of a point *in* X. This becomes apparent if one reverses the direction of the inclusion symbols in the definition. Then for an instance FL1 and FL2 read informally "$\mathfrak{f} \in X$" and "$\mathfrak{f} \notin \varnothing$", as it is expected when we think of $\mathfrak{f}$ as a point. The other conditions fit well into this scheme, too.

To construct filters it is often useful to start with a **filter base** $\mathfrak{b}$, i.e. a subset of $\mathfrak{P}(X)$ satisfying

FB1 $\mathfrak{b} \neq \varnothing$,

FB2 $\varnothing \notin \mathfrak{b}$,

FB3 $A, B \in \mathfrak{b}$ implies the existence of $C \subset A \cap B$ such that $C \in \mathfrak{b}$.

Then it is clear that

$$\overleftarrow{\mathfrak{b}} := \{A \subset X \mid \exists B \in \mathfrak{b} : B \subset A\}$$

is a filter.

A maximal filter with respect to inclusion as a subset of $\mathfrak{P}(X)$ is an **ultrafilter**. An ultrafilter $\mathfrak{u}$ is **principal** if it consists of all sets containing a fixed point $x \in X$, i.e. $\mathfrak{u} = \overline{\{\{x\}\}}$. In this manner we have a one-to-one correspondence between principal ultrafilters and actual points in X (which corroborates the picture of filters as generalized points.) The existence of non-principal ultrafilters is guaranteed by the axiom of choice. In particular the Ultrafilter Lemma states that every filter is contained in an ultrafilter. Apart from mere existence the important basic facts concerning ultrafilters are the following.

PROPOSITION 3.1 *Let $\mathfrak{f}$ be a filter in X. Then the following are equivalent.*

(1) *$\mathfrak{f}$ is an ultrafilter.*

(2) *For every subset $A \subset X$ either $A \in \mathfrak{f}$ or $X \setminus A \in \mathfrak{f}$*

(3) *If $A_1, \ldots, A_n$ are subsets of X such that $\bigcup_{i=1}^{n} A_i \in \mathfrak{f}$, then there is i such that $A_i \in \mathfrak{f}$.*

An important consequence is that an ultrafilter contains a finite set if and only if it is principal.

From now on the symbol I will mean a generic index set (usually infinite) and $\mathfrak{u}$ a non-principal ultrafilter in I.

When dealing with ultrafilters we introduce the following abbreviating notation. We say that a property P holds **$\mathfrak{u}$-almost everywhere** or for **$\mathfrak{u}$-almost all** i if the set $\{i \in I \,|\, P(i)\}$ is in $\mathfrak{u}$. We also write $P(i)\ [\mathfrak{u}]$ in this situation.

Let A_i be a family of algebraic structures where $i \in I$ and I is an arbitrary index set. For elements $a = (a_i)_{i\in I}$ and $b = (b_i)_{i\in I}$ in the direct product of the A_i we write $a \equiv_{\mathfrak{u}} b$ if $\{i \in I \,|\, a_i = b_i\} \in \mathfrak{u}$, i.e. $a_i = b_i\ [\mathfrak{u}]$. The filter properties readily imply that $\equiv_{\mathfrak{u}}$ is an equivalence relation. Then the ultraproduct of the A_i is defined as

$$\prod_{i\to\mathfrak{u}} A_i \coloneqq \left(\prod_{i\in I} A_i\right) \Big/ \equiv_{\mathfrak{u}}.$$

When the right index is understood we simply write $\prod_{\mathfrak{u}} A_i$. We shall use the notation $\boldsymbol{a}$ for the equivalence class in the ultraproduct with representative $a = (a_i)_{i\in I}$.

Let x_i be elements in a topological space X for all $i \in I$. If $x \in X$ has the property that for every neighborhood U of x the set $\{i \mid x_i \in U\}$ is contained in the filter $\mathfrak{f}$, then $(x_i)_{i\in I}$ **converges** to x along $\mathfrak{f}$. The following basic result applies.

PROPOSITION 3.2 *If X is a compact topological space and $\mathfrak{u}$ an ultrafilter, then every sequence $(x_i)_{i\in I}$ converges in X along $\mathfrak{u}$. If X is a Hausdorff space, then the limit point is unique.*

As a consequence in a compact Hausdorff space we can write

$$\lim_{\mathfrak{u}} x_i := \lim_{i\to\mathfrak{u}} x_i := x$$

when (x_i) converges to x along $\mathfrak{u}$. The unique point x is called the **ultralimit** of (x_i). We shall use the same notation in the context of non-compact spaces, for example $\mathbb{R}$. If an ultralimit does not exist therein we write $\lim_{\mathfrak{u}} x_i := \infty$.

Note that the ultraproduct construction is feasible with arbitrary filters, leading to so called **reduced products**. Nevertheless the further development relies on the following two results and hence requires ultrafilters.

PROPOSITION 3.3 *Let A_i be algebraic structures for all $i \in I$ and $J \subset I$. If $J \in \mathfrak{u}$, then*

$$\mathfrak{u}|J := \{K \cap J \mid K \in \mathfrak{u}\}$$

is a non-principal ultrafilter on J and the ultraproducts $\prod_{\mathfrak{u}} A_i$ and $\prod_{\mathfrak{u}|J} A_i$ are isomorphic.

P r o o f. The definition of ultraproducts readily implies that

$$\prod_{\mathfrak{u}} A_i \to \prod_{\mathfrak{u}|J} A_i, \qquad \boldsymbol{a} \mapsto (a_i)/\equiv_{\mathfrak{u}|J}$$

is a well defined isomorphism. □

THEOREM 3.4 (Łoś) *Let A_i, $i \in I$, be algebraic structures of the same signature. Then the ultraproduct $\prod_{\mathfrak{u}} A_i$ is an algebraic structure of the same signature. If P is a sentence in the language of the A_i which is expressible in first order logic, then P is true in $\prod_{\mathfrak{u}} A_i$ if and only if P is true $\mathfrak{u}$-almost everywhere.*

Proof. Confer Chapter V, proof of Theorem 2.9 in [9]. □

To get a grasp on these concepts an example seems reasonable. Consider for an instance groups G_i, $i \in I$, where some of the G_i are abelian and others are not. Let $A(G_i)$ express the property that G_i is abelian. Then first, Theorem 3.4 implies that the ultraproduct $\boldsymbol{G} := \prod_{\mathfrak{u}} G_i$ is a group. Since I is partitioned into I_1 and $I_2 = I \setminus I_1$, where $I_1 := \{i \in I \,|\, A(G_i)\}$, by Proposition 3.1 either $I_1 \in \mathfrak{u}$ or $I_2 \in \mathfrak{u}$. Hence by Proposition 3.3 $\boldsymbol{G}$ is isomorphic to $\boldsymbol{G}_j := \prod_{\mathfrak{u}|I_j} G_i$ where j equals 1 or 2. Again by Theorem 3.4 and since being abelian can be expressed in first order logic in the language of groups, $\boldsymbol{G}_1$ is abelian and $\boldsymbol{G}_2$ is not. Stated differently $\boldsymbol{G}$ is abelian if and only if $\mathfrak{u}$-almost all G_i are. Thus the decision which of the two possibilities applies for $\boldsymbol{G}$ is left to the ultrafilter.

§ 2 METRIC ULTRAPRODUCTS

Consider a collection of pseudometric spaces M_i, $i \in I$, the diameter of which is uniformly bounded (i. e. less than a constant K, independently of $i \in I$.) The limit

$$d(\boldsymbol{x}, \boldsymbol{y}) := \lim_{\mathfrak{u}} d(x_i, y_i)$$

defines a pseudometric on the ultraproduct $\boldsymbol{M} := \prod_{\mathfrak{u}} M_i$. As for all pseudometric spaces the relation $\boldsymbol{x} \sim_{\mathfrak{u}} \boldsymbol{y}$ if and only if $d(\boldsymbol{x}, \boldsymbol{y}) = 0$ defines an equivalence relation and the quotient space $\boldsymbol{M} / \sim_{\mathfrak{u}}$ is a metric space in a natural way. Following the notation in [4] we write $\left(\prod M_i\right)_{\mathfrak{u}}$ or $(\boldsymbol{M})_{\mathfrak{u}}$ for the quotient space, when the metric cannot be mistaken. The reader is encouraged to confer ibid. for the details. Note that we will write $\boldsymbol{x}$ not only for elements in $\boldsymbol{M}$ but also for equivalence classes in $(\boldsymbol{M})_{\mathfrak{u}}$. Since both constructions will not appear simultaneously, there should be no confusion.

In the case of groups G_i with invariant pseudo length function ℓ_i, we define

$$\ell(\boldsymbol{g}) := \lim_{\mathfrak{u}} \ell_i(g_i).$$

Then dividing by $\sim_{\mathfrak{u}}$ is the same as dividing by the normal subgroup

$$\boldsymbol{N} := \{\boldsymbol{g} \in \boldsymbol{G} \,|\, \ell(\boldsymbol{g}) = 0\}.$$

Hence the quotient group $G/N = (G)_{\mathrm{u}}$ is a group with invariant metric. We think of N as consisting of **infinitesimal** elements.

The appealing Example 2.1 in [37] shows that we are bound to use *invariant* pseudo length functions when studying metric ultraproducts of groups. In fact the length function ℓ obtained as an pointwise ultralimit cannot be expected to be invariant if the length functions ℓ_i are not invariant. Since we want N to be a normal subgroup, invariance of ℓ cannot be disregarded.

We provide one further example for future reference. Consider rings A_i with rank (or norm) function P_i. Then $(A)_{\mathrm{u}} := \left(\prod A_i\right)_{\mathrm{u}}$ is a ring with pseudo rank (or norm) function P defined by

$$P(a) := \lim_{\mathrm{u}} P_i(a_i).$$

Consider the ideal

$$N_P := \{a \in A \,|\, P(a) = 0\}.$$

Then $(A)_{\mathrm{u}} = (\prod_{\mathrm{u}} A_i)/N_P$ and as a corollary of Theorem 2.22 we deduce that the rank (or norm) group $((A)_{\mathrm{u}})_1^\times$ is isomorphic to $\left(\prod (A_i)_1^\times\right)_{\mathrm{u}}$.

At this point it becomes apparent why we insisted on the concept of asymptotic equivalence in Section 2. For two pseudo length functions to be locally asymptotically equivalent means to determine the same set of infinitesimal elements.

The effect on metric ultraproducts of groups is the following.

PROPOSITION 3.5 *Let ℓ_1 and ℓ_2 be locally asymptotically equivalent length functions in a class of groups $\{G_n \,|\, n \in \mathbb{N}\}$. Let u be a non-principal ultrafilter in I and $n\colon I \to \mathbb{N}$ such that $\lim_{\mathrm{u}} n_i = \infty$. In the ultraproduct $G := \prod_{i\to \mathrm{u}} G_{n_i}$ define*

$$\ell_j(g) := \lim_{\mathrm{u}} \ell_j(g_i)$$

and normal subgroups

$$N_j := \left\{g \in G \,\middle|\, \ell_j(g) = 0\right\}$$

for $j = 1,2$. Then the metric ultraproduct groups arising as quotients G/N_1 and G/N_2 are isomorphic.

Proof. Let ℓ_1 be bounded by ℓ_2 in radius δ with modulus c. If $\boldsymbol{g} \in \boldsymbol{N}_2$, then $\ell_2(g_i) < c^{-1}\varepsilon$ [u] for all $0 < \varepsilon < c^{-1}\delta$. Now there is $N \in \mathbb{N}$ such that $n_i \geq N$ implies $\ell_1(g) \leq c\ell_2(g)$ for all $g \in G_{n_i}$, and $\{i \mid n_i \geq N\} \in \mathrm{u}$, because $\lim_{\mathrm{u}} n_i = \infty$. Hence $\ell_1(g_i) < \varepsilon$ [u], which means $\boldsymbol{g} \in \boldsymbol{N}_1$. The inclusion $\boldsymbol{N}_1 \subset \boldsymbol{N}_2$ follows by symmetry. □

PROPOSITION 3.6 *Let G_i be a group and ℓ_i an invariant pseudo length function on G_i for all $i \in I$. We write $N_i := \{g \in G_i \mid \ell_i(g) = 0\}$. Then the groups G_i/N_i are naturally equipped with length functions $(\ell_i)_{G_i/N_i}$ and*

$$\left(\prod G_i\right)_{\mathrm{u}} \cong \left(\prod G_i/N_i\right)_{\mathrm{u}},$$

where the metric ultraproducts are formed using the respective length functions.

Proof. By Proposition 2.3 $(\ell_i)_{G_i/N_i}$ is a length function. If $\boldsymbol{g} \in \left(\prod G_i\right)_{\mathrm{u}}$ has a representative $(g_i)_{i\in I}$, then mapping $\boldsymbol{g}$ to the element in $\left(\prod G_i/N_i\right)_{\mathrm{u}}$ represented by $(g_iN_i)_{i\in I}$ is a well defined isomorphism. This follows simply from the definition of metric ultraproducts and the fact that $\ell_i(g) = (\ell_i)_{G_i/N_i}(gN_i)$ for all $i \in I$ and $g \in G_i$. □

The previous proposition can be applied for example to ultraproducts of general linear groups equipped with the Jordan length. Then we find that $\left(\prod \mathrm{GL}_{n_i}(K_i)\right)_{\mathrm{u}}$ is isomorphic to $\left(\prod \mathrm{PGL}_{n_i}(K_i)\right)_{\mathrm{u}}$, since $\ell_{\mathrm{J}}(g) = 0$ if and only if $g \in Z(\mathrm{GL}_{n_i}(K_i))$.

§ 3 GROUP APPROXIMATION

A good deal of modern mathematics consists of the study of mathematical objects of a certain type (e.g. groups) by analyzing morphisms into simpler objects. The usefulness of this approach cannot be contradicted at least since the introduction of Fourier analysis, and the whole concept of representation theory is an incarnation thereof. As it is useful the approach has its boundaries. To stick to the two above examples, Fourier analysis works at its best in the setting of abelian groups and classical representation theory experiences more and more drawbacks when moving from finite groups and Lie groups to locally compact groups and beyond. While we surely cannot expect *better* results by exchanging representation theory with something more general, we can still hope to find *some* results for *more* groups in doing so. The idea how to proceed is group approximation as explained in what follows.

Let $\mathcal{G}$ be a class of groups, each of which comes equipped with a pseudo length function ℓ. Then a group Γ is said to have the **$\mathcal{G}$-approximation property** if for all $g \in \Gamma \setminus \{1\}$ there is $\delta_g > 0$ such that for any $\varepsilon > 0$ and any finite subset $E \subset \Gamma$ there is a group $G \in \mathcal{G}$ and a mapping $\varphi \colon \Gamma \to G$ such that

AH1 $\ell(\varphi(1)) \leq \varepsilon$,

AH2 $\ell(\varphi(g)) \geq \delta_g$ for all $g \in E \setminus \{1\}$,

AH3 $\ell(\varphi(g)\varphi(h)\varphi(gh)^{-1}) \leq \varepsilon$ for all $g, h \in E$.

The kind of mapping in this definition is called **(E, ε)-homomorphism** or less explicit **almost homomorphism**. Note that almost homomorphisms depend not only on E and ε but also on the distribution of numbers δ_g, $g \in \Gamma$. Another method of approximating groups which is used often we will adress here as the **discrete $\mathcal{G}$-approximation property**. A group has this property if we replace the constants δ_g in the above definition by a common constant δ which does only depend on the whole of Γ. The **strong discrete $\mathcal{G}$-approximation property** demands $\ell(\varphi(g)) \geq \operatorname{diam}(G) - \varepsilon$ for all $g \in E \setminus \{1\}$ instead of AH2. Such φ is called a **strong** almost homomorphism. Of course this only makes sense if the groups $G \in \mathcal{G}$ have finite diameter.

Right after the definition the lack of strength in comparison with representation theory is obvious. Neither do we insist on the existence of morphisms into suitable groups, nor need our replacement – the almost homomorphism – be defined on the whole group, but only on a finite subset. The suggestive explanation referring to representation theory is only part of the historical motivation for group approximation. We shall follow the original idea and introduce on the way the two most prominent examples of group approximation, sofic and hyperlinear groups.

A group Γ is said to be **residually finite** if for every $g \in \Gamma \setminus \{1\}$ there is a homomorphism φ into a finite group such that $\varphi(g) \neq 1$. The notion of residual finiteness still follows the representation theory approach, where finite groups serve as simple objects to carry out investigation of the residually finite group. A somewhat different idea is amenability. A discrete group Γ is **amenable** if for every $\varepsilon > 0$ and every finite subset $E \subset \Gamma$ there exists an (E,ε)-**Følner set** F, i.e. a finite subset of Γ such that for all $g \in E$

$$|gF \triangle F| < \varepsilon |F|,$$

where $\triangle$ denotes the symmetric difference. Note that there is a plethora of definitions of amenability at hand. (According to [7], p. 48 there are $10^{10^{10}}$.) The approach using Følner sets is one of the more practical and we will use it on coming occasions.

In [23] Gromov sought a common generalization of amenability and residual finiteness. Weiss in [43] named the groups subject to this generalization sofic groups. Elek and Szabó gave the following characterization in terms of group approximation in [14]: A group is **sofic** if it has the $\mathcal{S}$-approximation property, where the Hamming length is used to measure distances in permutation groups.

Let Γ be a residually finite group. Then for any finite subset E of Γ and every $g \in E \setminus \{1\}$ there is a homomorphism φ_g into a finite group G_g such that $\varphi_g(g) \neq 1$. By Cayley's Theorem $\prod_{g\in E} G_g$ embeds into a symmetric group S_n by means of the permutation action on itself. Let

$$\varphi : x \mapsto (\varphi_g(x))_{g\in E} \in \prod_{g\in E} G_g \subset S_n.$$

Since φ is an injective homomorphism by construction, and non-trivial permutations in its range have no fixed points, φ is an (E,ε)-homomorphism for every $\varepsilon > 0$.

Now let Γ be amenable, $\varepsilon > 0$ and $E \subset \Gamma$ finite. Let F be an $(E, \frac{1}{2}\varepsilon)$-Følner set. Then for $g \in E^2$ we define $\varphi : \Gamma \to S_F$ by $\varphi_g(x) := gx$ if $gx \in F$ and extend φ_g to a bijection of F. Furthermore we let $\varphi_g = 1$ if $g \notin E^2$. We certainly have $\varphi_1 = 1$ and $\ell_{\mathrm{H}}(\varphi_g) \geq 1 - \frac{1}{2}\varepsilon$ for all $g \in E$. Moreover $\varphi_g\varphi_h(x) = \varphi_{gh}(x)$ if x is in the intersection $g^{-1}F \cap h^{-1}F \cap h^{-1}g^{-1}F \cap F$. This set has at least $(1-\varepsilon)|F|$ elements and hence $\ell_{\mathrm{H}}(\varphi_g\varphi_h\varphi_{gh}^{-1}) < \varepsilon$. Thus we know that indeed residually finite and amenable groups are sofic. (We shall henceforth usually write φ_g instead of $\varphi(g)$ if φ takes values in a group acting on some set.)

The second important example are hyperlinear groups. A group is **hyperlinear** if has the approximation property in the class of complex unitary groups $\mathrm{U}_n(\mathbb{C})$ with the length function ℓ_2, introduced in Section 2, § 4. In [15], Theorem 2 it is proved that every (countable) sofic group is hyperlinear. The proof of Theorem 3.3 in [37] actually shows that after embedding S_n as the subgroup of permutation matrices into $\mathrm{U}_n(\mathbb{C})$, $\ell_{\mathrm{H}}(\pi) \leq \frac{1}{2}\ell_2(P_\pi)^2$, where P_π is the permutation matrix corresponding to $\pi \in S_n$. Note that this is an even stronger condition than asymptotic boundedness. We will not treat hyperlinear groups in the following and refer the reader to [37] for further directions.

The main problem in the study of sofic groups is the question whether all groups are sofic. While the common opinion seems to expect a negative answer, concrete plans how to find a counterexample are still elusive. A first step towards a better understanding of sofic groups could be investigation of other instances of group approximation. To pincer the problematic sofic groups one approach is to investigate group approximation in a more general setting to find a non-sofic group, or to show that more general notions of group approximation can be reduced to approximation in symmetric groups. The hyperlinear groups have still not proven to be useful for one of these goals. Another approach is to use more restrictive classes of groups for approximation to explore techniques for finding counterexamples to approximation properties. In [41] Thom showed that Higman's group (introduced in [26]) cannot be approximated using groups with **commutator contractive** length functions, i.e. length functions satisfying $\ell([a,b]) \leq 4\ell(a)\ell(b)$. This seems to be the only natural example in this direction so far.

The fact that all known groups are either sofic or of undecided status entails that theorems known to be true for sofic groups hold for nearly as many groups as we can imagine today. Results proven for sofic groups include Gottschalk's Surjunctivity Conjecture (proved in [23]), Kaplansky's Direct Finiteness Conjecture (proved in [14] and Connes' Embedding Conjecture. These questions are adressed in more detail in [37]. Their positive answer for sofic groups are examples of the usefulness of the very general idea of group approximation as suggested at the beginning of this paragraph.

We proceed by exhibiting the connection between group approximation and metric ultraproducts. The following fundamental theorem is a generalization of Theorem 1 in [15]. Confer also [41], Proposition 1.8.

THEOREM 3.7 *Let Γ be a group. Then Γ has the $\mathcal{G}$-approximation property if and only if there is a suitable index set I and an ultrafilter $\mathfrak{u}$ in I such that Γ can be embedded into a metric ultraproduct $(\boldsymbol{G})_{\mathfrak{u}} := \big(\prod_{i\in I} G_i\big)_{\mathfrak{u}}$ with groups $G_i \in \mathcal{G}$. The set I can be chosen to have cardinality not exceeding the cardinality of Γ.*

P r o o f. The case of Γ being finite is trivial. Hence assume that Γ is infinite with the $\mathcal{G}$-approximation property. We consider pairs (E,ε) of finite subsets E of Γ and rational numbers ε in the interval $]0,1[$. Call the set of these I and define subsets

$$I_{H,\delta} := \{(E,\varepsilon)\in I \mid H \subset E, \varepsilon \le \delta\}.$$

Then the set of all $I_{H,\delta}$ forms a filter base, which is contained in an ultrafilter $\mathfrak{u}$. For every $(E,\varepsilon)\in I$ choose an (E,ε)-homomorphism $\varphi_{E,\varepsilon}$ into a group $G_{E,\varepsilon}$. Then we can map Γ into the direct product of the $G_{E,\varepsilon}$ by $\varphi'\colon g \mapsto (\varphi_{E,\varepsilon}(g))_{E,\varepsilon}$. By composing with the quotient homomorphism $\pi\colon \prod G_{E,\varepsilon} \to (\boldsymbol{G})_{\mathfrak{u}}$, we obtain a mapping $\varphi := \pi\circ\varphi'\colon \Gamma \to (\boldsymbol{G})_{\mathfrak{u}}$. Note that $\lim_{(E,\varepsilon)\to\mathfrak{u}} \varepsilon = 0$ by the choice of $\mathfrak{u}$. Let g,h be arbitrary in Γ. We first deduce

$$\ell(\varphi(g)\varphi(h)\varphi(gh)^{-1}) = \lim_{\mathfrak{u}} \ell(\varphi_{E,\varepsilon}(g)\varphi_{E,\varepsilon}(h)\varphi_{E,\varepsilon}(gh)^{-1}) \le \lim_{\mathfrak{u}} \varepsilon = 0,$$

whence $\varphi(gh) = \varphi(g)\varphi(h)$ and φ is a homomorphism. Further

$$\ell(\varphi(g)) = \lim_{\mathfrak{u}} \ell(\varphi_{E,\varepsilon}(g)) \ge \lim_{\mathfrak{u}} \delta_g = \delta_g$$

for $g \ne 1$. Therefore $\varphi(g)\ne 1$ and φ is injective. Note that by definition of I its cardinality equals the cardinality of Γ.

To show the other implication suppose we are in control of I, $\mathfrak{u}$ and an injective homomorphism $\varphi\colon \Gamma \to (\boldsymbol{G})_{\mathfrak{u}}$. We define $\delta_g := \frac{1}{2}\ell(\varphi(g))$, where ℓ is the ultralimit length function on $(\boldsymbol{G})_{\mathfrak{u}}$. Let $E \subset \Gamma$ be a finite subset and $\varepsilon > 0$. Then there is a mapping $\psi\colon \Gamma \to \pi^{-1}(\big(\prod G_i\big)_{\mathfrak{u}})$ such that $\pi\circ\psi = \varphi$. Let ψ_i be the projection of ψ onto G_i. Then

$$\lim_{\mathfrak{u}} \ell(\psi_i(g)) = \ell(\varphi(g)) = 2\delta_g$$

for fixed $g\in E$, and therefore $\ell(\psi_i(g)) \ge \delta_g$ $[\mathfrak{u}]$. Likewise

$$\ell(\psi_i(g)\psi_i(h)\psi_i(gh)^{-1}) < \varepsilon \;\; [\mathfrak{u}]$$

for $g, h \in E$. Hence by finiteness of E, the set of all i such that ψ_i is an (E,ε)-homomorphism, is contained in $\mathfrak{u}$. Thus it is not empty and we may pick a convenient index i, providing us with a group G_i in $\mathcal{G}$ and an (E,ε)-homomorphism ψ_i. □

By closer inspection of the proof, we obtain the following result.

COROLLARY 3.8 *A group has the discrete $\mathcal{G}$-approximation property if and only if it embeds into $(\mathbf{G})_{\mathfrak{u}}$ as a discrete subgroup.*

Note that the very definition of $\mathcal{G}$-approximation immediately implies that a group Γ has the $\mathcal{G}$-approximation property if and only if every finitely generated subgroup does. Hence it often suffices to study countable groups with the $\mathcal{G}$-approximation property. The same is true for the discrete and strong discrete approximation property.

To make further investigation a bit more pleasant, we relax the conditions characterizing almost homomorphisms.

PROPOSITION 3.9 *Condition AH1 is not necessary in the definition of almost homomorphisms, provided $1 \in E$. If Γ has the $\mathcal{G}$-approximation property, then for all finite $E \subset \Gamma$ and $\varepsilon > 0$ there is an (E,ε)-homomorphism φ such that $\varphi(1) = 1$. Furthermore for $g \in E$ the length $\ell(\varphi(g))$ is as large as we can expect of any almost homomorphism $\Gamma \to G \in \mathcal{G}$.*

P r o o f. Let $E \subset \Gamma$ be finite and $\varepsilon > 0$, and assume $1 \in E$. Let $\psi\colon \Gamma \to G$ be a mapping satisfying AH2 and AH3. Then

$$\ell(\psi(1)) = \ell(\psi(1) \cdot \psi(1)\psi(1 \cdot 1)^{-1}) \leq \varepsilon$$

follows.

Now assume ψ is an $(E, \frac{1}{2}\varepsilon)$-homomorphism. We define φ to take the same values as ψ does, except $\varphi(1) := 1$. It suffices to show $\ell(\varphi(g)\varphi(h)\varphi(gh)^{-1}) \leq \varepsilon$ for $g, h \in E \cup \{1\}$ and $gh \in E^2 \cup \{1\}$. The interesting cases are without loss of generality $g = h^{-1}$, $g \neq 1$, and $g \neq 1$, $h = 1$. In the first case

$$\begin{aligned}
\ell(\varphi(g)\varphi(g^{-1})) &= \ell(\psi(g)\psi(g^{-1})\psi(1)^{-1}\psi(1)) \\
&\leq \ell(\psi(g)\psi(g^{-1})\psi(gg^{-1})^{-1}) + \ell(\psi(1)) \\
&\leq \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon = \varepsilon.
\end{aligned}$$

In the second case we are left to check $\ell(\varphi(g)\varphi(g)^{-1}) \leq \varepsilon$. This is true, because

$$\begin{aligned}\ell(\varphi(g)\varphi(g)^{-1}) &= \ell(\psi(g)\psi(g^{-1})\psi(1)^{-1}\psi(1)\psi(g^{-1})^{-1}\psi(g)^{-1}) \\ &\leq \ell(\psi(g)\psi(g^{-1})\psi(1)^{-1}) + \ell(\psi(1)\psi(g^{-1})^{-1}\psi(g)^{-1}) \\ &\leq \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon = \varepsilon,\end{aligned}$$

and the proof is complete. □

As a motivation for a sensible generalization of soficity, we give the following result.

Theorem 3.10 *A group Γ is sofic if and only if it has the $\mathcal{S}$-approximation property with respect to the conjugacy length.*

Proof. The claim for finite groups follows from Cayley's Theorem, whence we assume Γ to be infinite. By Theorem 3.7 (as was first proved by Elek and Szabó in [15], Theorem 1) Γ is sofic if and only if it embeds into a metric ultraproduct

$$(\mathcal{S})_{\mathrm{u}} = \left(\prod S_{n_i}\right)_{\mathrm{u}}$$

with respect to the Hamming length. Since Γ is infinite and an ultraproduct of groups of bounded cardinality is finite, we necessarily have $\lim_{\mathrm{u}} n_i = \infty$. According to Theorem 2.14 and Proposition 3.5 $(\mathcal{S})_{\mathrm{u}}$ is isomorphic to the metric ultraproduct of groups S_{n_i} with respect to the conjugacy length. Another application of Theorem 3.7 ensures that this is equivalent to Γ having the $\mathcal{S}$-approximation property with respect to the conjugacy length. □

Since the conjugacy length can be defined generically on every finite group, it offers the chance of a broad generalization of the notion of soficity. Of course on an even wider scope one could investigate groups which can be approximated with finite groups endowed with *any* length function. This idea first appears in [22] and there groups with the approximation property in the class of finite groups with no restriction on the length function were named **weakly sofic**. However, there is still no reasonable means to get a hold on all length functions on finite groups. Instead we want to motivate the study of approximation using matrix groups with the Jordan length. The first practical reason is that we already know from Section 2, § 5 that the conjugacy length and the Jordan length on general linear groups over finite fields are asymptotically equivalent. Thus approximating with the Jordan length is in accordance with the universal conjugacy length approach. Moreover matrix groups can be conveniently handled using the tools of linear algebra. By personal communication with Nikolay Nikolov the author was informed that every simple weakly sofic group is sofic or can be approximated in general linear groups over finite fields with the Jordan length. (This result can be deduced from Theorem 1.7 in[34].) Hence this kind of approximation seems to be a natural and useful approach.

Let $\mathcal{GL}$ denote the class of all general linear groups over arbitrary fields. We shall call groups with the $\mathcal{GL}$-approximation property using the rank length **linearly sofic** groups. If a group has the $\mathcal{GL}(K)$-approximation property for a fixed field K we will call it **K-sofic**. When instead approximation is done using the Jordan length we will speak of **projectively linearly sofic** and **projectively K-sofic** groups, respectively.

§ 4 AMPLIFICATION PROPERTIES

In the definition of group approximation whether a group Γ has the *(strong) discrete* approximation property a priori depends on Γ. In certain classes $\mathcal{G}$ we can enforce the (strong) discrete $\mathcal{G}$-approximation property for every group having the $\mathcal{G}$-approximation property: A class of groups $\mathcal{G}$ has the **amplification property** if there exists $\delta > 0$ such that for any group Γ and for all $\varepsilon > 0$ there exists $\varepsilon' > 0$ such that the following holds: Whenever $\varphi\colon \Gamma \to G$ is an (E,ε')-homomorphism into $G \in \mathcal{G}$, then there is $H \in \mathcal{G}$ and a mapping $\iota\colon G \to H$ such that $\iota \circ \varphi\colon \Gamma \to H$ is an (E,ε)-homomorphism with the additional property that $\ell(\iota \circ \varphi(g)) \geq \delta$ for all $g \in E \setminus \{1\}$. If there exist H and ι such that $\iota \circ \varphi$ satisfies $\ell(\iota \circ \varphi(g)) \geq \operatorname{diam}(H) - \varepsilon$ we say that $\mathcal{G}$ has the **strong amplification property**. (Note that δ is no longer needed in the second definition.)

Now the next proposition follows directly.

PROPOSITION 3.11 *If $\mathcal{G}$ has the (strong) amplification property, then any group Γ has the (strong) discrete $\mathcal{G}$-approximation property if and only if it has the $\mathcal{G}$-approximation property.*

The maybe best known example of amplification are the symmetric groups. The proof appears e. g. in [15], proof of Theorem 1 or in [37], proof of Theorem 3.5.

PROPOSITION 3.12 *The class $\mathcal{S}$ of symmetric groups with the Hamming length has the strong amplification property.*

Consider the following statements about approximation of subgroups, inverse limits and direct products of groups.

PROPOSITION 3.13 *Let $\mathcal{G}$ be a class of groups with invariant pseudo length functions. Then the class of groups with the $\mathcal{G}$-approximation property is closed under taking subgroups and inverse limits. The same is true for the discrete and strong discrete $\mathcal{G}$-approximation property.*

Proof. The statement concerning subgroups is obvious.

Assume that Γ is the inverse limit of groups Γ_i with projection morphisms $\pi_i^j : \Gamma_i \to \Gamma_j$ and $\pi_i : \Gamma \to \Gamma_i$, where i, j are from a directed set I. Then Γ can be identified with the set of vectors $(g_i)_{i\in I}$ in $\prod_{i\in I} \Gamma_i$ such that $\pi_i^j(g_i) = g_j$ if $j \leq i$. If E is a finite subset of Γ there is an index i such that $\pi_i(g) \neq \pi_i(h)$ for all $g, h \in E^2$. By assumption for any $\varepsilon > 0$ there is a $(\pi_i(E), \varepsilon)$-homomorphism $\varphi : \Gamma_i \to G$ for some $G \in \mathcal{G}$. By the choice of i, $\varphi \circ \pi_i$ is well defined on E^2. Then some arbitrary extension of $\varphi \circ \pi_i$ to the rest of Γ is automatically an (E, ε)-homomorphism. The same argument works for the stronger $\mathcal{G}$-approximation properties in the claim. $\square$

PROPOSITION 3.14 *Suppose that $\mathcal{G}$ has the following property: For every finite direct product $G_1 \times \ldots \times G_k$ of groups $G_i \in \mathcal{G}$ there are weights $w_1, \ldots, w_k \in [0, \infty[$, a group $H \in \mathcal{G}$ and an isometric embedding $G_1 \times \ldots \times G_k \to H$, where we use the pseudo length function*

$$\ell((g_1, \ldots, g_k)) := \frac{\sum_{i=1}^k w_i \ell(g_i)}{\sum_{i=1}^k w_i}$$

on $G_1 \times \ldots \times G_k$. Then the class of groups with the $\mathcal{G}$-approximation property is closed under taking direct products. The same holds for the discrete and strong discrete $\mathcal{G}$-approximation property.

Proof. Because being approximated with groups in $\mathcal{G}$ is a local property, it clearly suffices to consider finitary direct products. Now the additional assumption on $\mathcal{G}$ implies the claim immediately. The argument also works for the discrete ond strong discrete approximation property. □

Certainly statements as above for other group theoretical constructions would be a great thing to have. Unfortunately under very general assumptions very little can be done. An example where the amplification property is needed is the following.

PROPOSITION 3.15 *If $\mathcal{G}$ has the amplification property, then the class of groups with the $\mathcal{G}$-approximation property is closed under taking direct limits.*

Proof. Let Γ be a direct limit of groups Γ_i, which can be approximated in $\mathcal{G}$. Let $\varepsilon > 0$ and $E \subset \Gamma$ be a finite subset. Because E is finite, eventually $E^2 \subset \Gamma_i$ holds. Thus we find an (E, ε)-homorphism $\varphi \colon \Gamma_i \to G$, where $G \in \mathcal{G}$, which can be extended arbitrarily to an almost homomorphism defined on Γ. Then $\ell(\varphi(g)) \geq \delta_{g,i} > 0$. Since a priori $\delta_{g,i}$ depends on i, we need the amplification property to ensure $\delta_{g,i} \geq \delta_g$ for constants δ_g not depending on i. □

In [3] Arzhantseva and Păunescu showed that $\{\mathrm{GL}_n(\mathbb{C}) \mid n \in \mathbb{N}\}$ with the rank length has the amplification property. In fact the proof does not use particular properties of the complex numbers apart from characteristic zero and it could be generalized to arbitrary fields of characteristic zero. We will modify the method of proof from [3], Section 5, to work in any characteristic and to show the amplification property also when working with the Jordan length. Note that *linear sofic groups* in [3] are $\mathbb{C}$-sofic groups in our sense.

Let K be an algebraically closed field. Then the Jordan decomposition as explained in Section 2, § 5 for matrices in $\mathrm{GL}_n(K)$ exists. ([5], Chapter I, § 4 offers a detailed treatment.) In particular in algebraically closed fields every square matrix is conjugate to a matrix in Jordan normal form.

Over any field we use the notation $J(\alpha, s)$ for $s \times s$-Jordan matrices with eigenvalue α. Every matrix $A \in \mathrm{GL}_n(K)$ is in $\mathrm{GL}_n(\overline{K})$ conjugate to a matrix A' in Jordan normal form, where $\overline{K}$ is the algebraic closure of K. We write $\iota_\alpha(A)$ for the number of Jordan blocks $J(\alpha,1)$ of A' divided by n. Furthermore let

$$\iota(A) := \sup_{\alpha\in K^\times} \iota_\alpha(A).$$

(That is $\iota(A) = \frac{m_{11}}{n}$ in the notation of Section 2, § 5.)

PROPOSITION 3.16 *Let K be a field and A in $\mathrm{GL}_n(K)$. Then*

$$\tfrac{1}{2}(1-\iota_1(A)) \le \ell_{\mathrm{r}}(A) \le 1-\iota_1(A), \qquad \tfrac{1}{2}(1-\iota(A)) \le \ell_{\mathrm{J}}(A) \le 1-\iota(A).$$

Proof. Let A' be a matrix in Jordan normal form conjugate to A in $\mathrm{GL}_n(\overline{K})$. If this matrix has k Jordan blocks $J(1,1)$, then

$$\ell_{\mathrm{r}}(A) = \frac{\mathrm{rk}(1-A)}{n} \le \frac{n-k}{n} = 1-\iota_1(A).$$

Since the remaining Jordan blocks do not have eigenvalues equal to 1 or are of size strictly larger than 1, also

$$\ell_{\mathrm{r}}(A) = \frac{\mathrm{rk}(1-A)}{n} \ge \frac{n-k}{2n} = \tfrac{1}{2}(1-\iota_1(A)).$$

The claimed inequalities for $\iota(A)$ and $\ell_{\mathrm{J}}(A)$ follow from the above result, since

$$\ell_{\mathrm{J}}(A) = \inf_{\alpha\in K^\times} \ell_{\mathrm{r}}(\alpha A)$$

and

$$\inf_{\alpha\in K^\times}(1-\iota_1(\alpha A)) = 1 - \sup_{\alpha\in K^\times}\iota_\alpha(A) = 1-\iota(A).$$

□

THEOREM 3.17 *Let K be an algebraically closed field and $J(\alpha,s)$, $J(\beta,t)$ two Jordan matrices with eigenvalues $\alpha,\beta\in K^\times$ and $s\le t$. Then the Jordan normal form of $J(\alpha,s)\otimes J(\beta,t)$ has s Jordan blocks.*

Proof. In characteristic 0 the claim follows from Corollary 2.2.11 in [27], in positive characteristic from Theorem 2.2.2, ibid. □

LEMMA 3.18 *Let A be a matrix in $\mathrm{GL}_n(K)$ and α an eigenvalues of A. Suppose that in a matrix A' in Jordan normal form, obtained from A over $\overline{K}$, the Jordan block corresponding to α has size s. If the extension $K(\alpha)/K$ is inseparable, then $s = p^k$, where p is the characteristic of K and $k > 0$.*

P r o o f. First of all we note that for $K(\alpha)/K$ to be inseparable K necessarily has to be of positive characteristic. Let f be the minimal polynomial of α over K. Let $\alpha_1,\ldots,\alpha_m$ be the roots of f, where $\alpha_i=\alpha_j$ if and only if $i=j$. By [30], Chapter V, Proposition 6.1

$$f=(x-\alpha_1)^{p^k}\cdot\ldots\cdot(x-\alpha_m)^{p^k}$$

for a natural number k, since f is inseparable. Moreover f divides the minimal polynomial μ_A of A, because every root of f is a root of μ_A. The remaining factor g such that $\mu_A=f\cdot g$ has only roots different from the roots of f. Hence the multiplicity of α as a root of μ_A is p^k. Therefore $\dim\ker(A-\alpha)^j>\dim\ker(A-\alpha)^{j-1}$ if and only if $j\in\{1,\ldots,p^k\}$. This means that the Jordan block corresponding to α has size p^k. □

LEMMA 3.19 *Let $x'\geq x\geq 0$, $y'\geq y\geq 0$ be real numbers. Then*

$$x'y+y'x\leq x'y'+xy.$$

P r o o f. We calculate

$$\begin{aligned}
2(x'y+y'x)&=x'(y+y'-y')+(x'+x-x)y+y'(x+x'-x')+(y'+y-y)x\\
&=x'y'+x'(y-y')+xy+(x'-x)y\\
&\qquad+y'x'+y'(x-x')+yx+(y'-y)x\\
&=2x'y'+2xy+(x'-x)(y-y')+(y'-y)(x-x')\\
&\leq 2(x'y'+xy)
\end{aligned}$$

to complete the proof. □

LEMMA 3.20 *Let K be a field and $A\in\mathrm{GL}_n(K)$, $B\in\mathrm{GL}_m(K)$. Then*

$$\iota(A\otimes B)\leq\iota(A)\iota(B)+(1-\iota(A))(1-\iota(B)).$$

If $\iota(A)\leq\frac{1}{2}$ and $\iota(B)\leq\frac{1}{2}$, then $\iota(A\otimes B)\leq\frac{1}{2}$.

P r o o f. We work with A embedded in $\mathrm{GL}_n(\overline{K})$ and B in $\mathrm{GL}_m(\overline{K})$. Let A', B' be matrices in Jordan normal form corresponding to A and B, respectively. Then it is clear that the Jordan normal form of $A \otimes B$ is the same as the one of $A' \otimes B'$, since using conjugation to compute the Jordan normal form commutes with the tensor product. To obtain the Jordan normal form of $A' \otimes B'$ it is clearly sufficient to compute the Jordan normal forms of $J(\alpha, s) \otimes J(\beta, t)$ for all combinations of Jordan blocks $J(\alpha, s)$ of A' and $J(\beta, t)$ of B'.

Since $J(\alpha, 1) \otimes J(\beta, t)$ equals $\alpha J(\beta, t)$, we know that on the one hand two Jordan blocks of size 1 yield a Jordan block of size 1 in the Jordan normal form of $A \otimes B$. On the other hand a Jordan block of size 1 and a larger one cannot produce a Jordan block of size 1. Moreover two Jordan blocks $J(\alpha, s)$ and $J(\beta, t)$, where $1 < s \leq t$, can be responsible for at most $s - 1$ Jordan blocks of size 1, by Theorem 3.17. Assume α is an eigenvalue of A or B such that $K(\alpha)/K$ is inseparable. Then by Lemma 3.18 α corresponds to a Jordan block of size larger than or equal to the characteristic of K, in particular strictly larger than 1. Denote the separable closure of K by $\overline{K}_s$ and let $\varkappa(A) := \sum_{\alpha \in \overline{K}_s^\times} \iota_\alpha(A)$. If $\gamma \in K^\times$ such that $\iota_\gamma(A \otimes B) = \iota(A \otimes B)$ we can deduce

$$\iota(A \otimes B) \leq \sum_{\alpha \in K^\times} \iota_\alpha(A) \iota_{\alpha^{-1}\gamma}(B) + \sum_{\alpha \in \overline{K}_s^\times \setminus K^\times} \iota_\alpha(A) \iota_{\alpha^{-1}\gamma}(B) + \tfrac{1}{2}(1 - \varkappa(A))(1 - \varkappa(B)).$$

Here the splitting in sums over $K^\times$ and $\overline{K}_s^\times \setminus K^\times$ is possible, because $K^\times$ is a subgroup of $\overline{K}_s^\times$.

Let λ, δ be in $K^\times$ such that $\iota_\lambda(A) = \iota(A)$ and $\iota_\delta(B) = \iota(B)$. Then

$$\begin{aligned}
\sum_{\alpha \in K^\times} \iota_\alpha(A) \iota_{\alpha^{-1}\gamma}(B) &= \iota_\lambda(A) \iota_{\lambda^{-1}\gamma}(B) + \iota_{\delta^{-1}\gamma}(A) \iota_\delta(B) + \sum_{\lambda, \delta^{-1}\gamma \neq \alpha \in K^\times} \iota_\alpha(A) \iota_{\alpha^{-1}\gamma}(B) \\
&\leq \iota_\lambda(A) \iota_\delta(B) + \iota_{\delta^{-1}\gamma}(A) \iota_{\lambda^{-1}\gamma}(B) + \sum_{\lambda, \delta^{-1}\gamma \neq \alpha \in K^\times} \iota_\alpha(A) \iota_{\alpha^{-1}\gamma}(B) \\
&\leq \iota_\lambda(A) \iota_\delta(B) + \sum_{\lambda \neq \alpha \in K^\times} \iota_\alpha(A) \sum_{\delta \neq \beta \in K^\times} \iota_\beta(B),
\end{aligned}$$

where we used Lemma 3.19. By the choice of γ, λ and δ, and the preceding estimate of $\iota(A \otimes B)$ we arrive at

$$\iota(A \otimes B) \leq \iota(A)\iota(B) + (1 - \iota(A))(1 - \iota(B)),$$

which proves the first claim.

Now assume $\iota(A), \iota(B) \leq \frac{1}{2}$. If the eigenvalues of A in K are λ_i such that $\iota_{\lambda_1}(A) \geq \iota_{\lambda_2}(A) \geq \dots$ and the eigenvalues of B in K are δ_i such that $\iota_{\delta_1}(B) \geq \iota_{\delta_2}(B) \geq \dots$, then we can proceed inductively from

$$\sum_{\alpha \in K^\times} \iota_\alpha(A)\iota_{\alpha^{-1}\gamma}(B) \leq \iota_{\lambda_1}(A)\iota_{\delta_1}(B) + \iota_{\delta_1^{-1}\gamma}(A)\iota_{\lambda_1^{-1}\gamma}(B) + \sum_{\lambda_1, \delta_1^{-1}\gamma \neq \alpha \in K^\times} \iota_\alpha(A)\iota_{\alpha^{-1}\gamma}(B)$$

to obtain

$$\sum_{\alpha \in K^\times} \iota_\alpha(A)\iota_{\alpha^{-1}\gamma}(B) \leq \sum_i \iota_{\lambda_i}(A)\iota_{\delta_i}(B).$$

If $K(\alpha)/K$ is separable, then α has at least one Galois conjugate eigenvalue. This implies

$$\sum_{\alpha \in \overline{K}_s^\times \setminus K^\times} \iota_\alpha(A)\iota_{\alpha^{-1}\gamma}(B) \leq \tfrac{1}{2} \sum_{\alpha \in \overline{K}_s^\times \setminus K^\times} \iota_\alpha(A) \sum_{\beta \in \overline{K}_s^\times \setminus K^\times} \iota_\beta(B).$$

Combining the different estimates proves $\iota(A \otimes B) \leq \frac{1}{2}$. □

PROPOSITION 3.21 ([3], Proposition 5.3) *Let $f : [\frac{1}{2}, 1] \to [\frac{1}{2}, 1]$ be defined by*

$$f(x) := x^2 + (1-x)^2.$$

Then f is a strictly monotone increasing bijection and $x \in [\frac{1}{2}, 1[$ implies

$$\lim_{m \to \infty} f^m(x) = \tfrac{1}{2}.$$

LEMMA 3.22 *Let $A \in \mathrm{GL}_n(K)$, $B \in \mathrm{GL}_m(K)$ and $\alpha \in K^\times$. Then*

(1) $\ell_{\mathrm{r}}(A \oplus B) = \frac{n}{n+m}\ell_{\mathrm{r}}(A) + \frac{m}{n+m}\ell_{\mathrm{r}}(B)$,

(2) $\ell_{\mathrm{r}}(A \otimes B) \leq \ell_{\mathrm{r}}(A) + \ell_{\mathrm{r}}(B)$,

(3) $\ell_{\mathrm{J}}(\alpha \cdot A \oplus B) \leq \min\{\frac{n}{n+m} + \frac{m}{n+m}\ell_{\mathrm{J}}(B), \frac{n}{n+m}\ell_{\mathrm{J}}(A) + \frac{m}{n+m}\}$ *and* $\ell_{\mathrm{J}}((\alpha A) \oplus B) = \frac{n}{n+m}\ell_{\mathrm{J}}(A) + \frac{m}{n+m}\ell_{\mathrm{J}}(B)$ *if* $\ell_{\mathrm{J}}(A) = \ell_{\mathrm{r}}(\alpha\beta A)$ *and* $\ell_{\mathrm{J}}(B) = \ell_{\mathrm{r}}(\beta B)$ *for some* $\beta \in K^\times$,

(4) $\ell_{\mathrm{J}}(\alpha \cdot A \otimes B) \leq \ell_{\mathrm{J}}(A) + \ell_{\mathrm{J}}(B)$.

P r o o f. Equation (1) follows from

$$\mathrm{rk}(1-A\oplus B)=\mathrm{rk}(1-A)+\mathrm{rk}(1-B).$$

The matrix $A\otimes B$ acts on K^{nm} by $A\otimes B(v\otimes w)=A(v)\otimes B(w)$ for all $v\in K^n$, $w\in K^m$ and linear extension. Therefore $A(v)=v$ and $B(w)=w$ implies $A\otimes B(v\otimes w)=v\otimes w$, whence $\dim\ker(1-A\otimes B)\geq \dim\ker(1-A)\cdot\dim\ker(1-B)$. Let $n_0:=\dim\ker(1-A)$ and $m_0:=\dim\ker(1-B)$. Then

$$\begin{aligned}
2\frac{nm-n_0m_0}{nm}&=\frac{nm-nm_0+nm_0-n_0m_0}{nm}+\frac{nm-mn_0+mn_0-n_0m_0}{nm}\\
&=\frac{m-m_0}{m}+\frac{m_0(n-n_0)}{nm}+\frac{n-n_0}{n}+\frac{n_0(m-m_0)}{nm}\\
&=\left(1+\frac{m_0}{m}\right)\frac{n-n_0}{n}+\left(1+\frac{n_0}{n}\right)\frac{m-m_0}{m}\\
&\leq 2\frac{n-n_0}{n}+2\frac{m-m_0}{m}
\end{aligned}$$

implies $\ell_{\mathrm{r}}(A\otimes B)\leq \ell_{\mathrm{r}}(A)+\ell_{\mathrm{r}}(B)$.

There are $\beta_1,\beta_2\in K^\times$ such that $\ell_{\mathrm{J}}(A)=\ell_{\mathrm{r}}(\alpha\beta_1A)$ and $\ell_{\mathrm{J}}(B)=\ell_{\mathrm{r}}(\beta_2B)$. If $\beta_1=\beta_2$, then $\ell_{\mathrm{J}}((\alpha A)\oplus B)=\frac{n}{n+m}\ell_{\mathrm{J}}(A)+\frac{m}{n+m}\ell_{\mathrm{J}}(B)$ follows from (1) and invariance of ℓ_{J} under scalar multiplication. Otherwise by the definition of ℓ_{J} as an infimum only the inequality in (3) holds.

At last (4) follows from (2) by linearity of the tensor product and invariance of ℓ_{J} under scalar multiplication. □

Let A be a matrix in $\mathrm{GL}_n(K)$. We write

$$A^{\otimes k}:=A\otimes\ldots\otimes A,$$

where A appears k times on the right side.

The construction in the next theorem is essentially from [3], Theorem 5.10. The proof is modified, though, to fit our treatment using almost homomorphisms, to work in arbitrary characteristic and when approximating with the Jordan length.

THEOREM 3.23 *Let K be a field. The class of groups $\{\mathrm{GL}_n(K) \mid n \in \mathbb{N}\}$ with rank length or Jordan length has the amplification property. In particular, for every projectively K-sofic group Γ, finite subset $E \subset \Gamma$ and $\varepsilon > 0$ there exists an (E,ε)-homomorphism φ satisfying $\ell_{\mathrm{J}}(\varphi(g)) > \frac{1}{4} - \varepsilon$ for all $g \in E$. In the case of Γ being K-sofic we can achieve the analogous estimate $\ell_{\mathrm{r}}(\varphi(g)) > \frac{1}{8} - \varepsilon$.*

Proof. We will first treat the case of approximation with the Jordan length. Consider any projectively K-sofic group Γ, $\varepsilon > 0$ and a finite subset $E \subset \Gamma$, and let $\delta := \min_{g\in E} \delta_g$. By Proposition 3.21 there is a natural number m such that $f^m(1-\delta) \leq \frac{1}{2} + 2\varepsilon$, where $f(x) := x^2 + (1-x)^2$. Choose $\varepsilon' > 0$ such that $\varepsilon' < 2^{-m}\varepsilon$. Then there exists an (E,ε')-homomorphism $\varphi\colon \Gamma \to \mathrm{GL}_n(K)$, where $\ell_{\mathrm{J}}(\varphi(g)) \geq \delta$ for all $g \in E$ and we can assume $\varphi(1) = 1$. Recursively define

$$\varphi_1(g) := \varphi(g), \qquad \varphi_{k+1}(g) := \varphi_k(g) \otimes \varphi_k(g).$$

We shall prove that $\varphi_m\colon \Gamma \to \mathrm{GL}_{n^{2^m}}(K)$ is an (E,ε)-homomorphism satisfying $\ell_{\mathrm{J}}(\varphi(g)) \geq \frac{1}{4} - \varepsilon$ for all $g \in E$. If $g \in E$, then $\iota(\varphi(g)) \leq 1 - \ell_{\mathrm{J}}(\varphi(g)) \leq 1 - \delta$. Therefore $\iota(\varphi_k(g)) \leq f^k(\iota(\varphi(g)))$ as long as $\iota(\varphi_{k-1}(g)) \geq \frac{1}{2}$. If $\iota(\varphi_{k-1}(g)) < \frac{1}{2}$ for one $k \leq m$, then $\iota(\varphi_k(g)) < f(\iota(\varphi_{k-1}(g))) < \frac{1}{2}$ by Lemma 3.20. Otherwise by the choice of m still $\iota(\varphi_m(g)) \leq \frac{1}{2} + 2\varepsilon$. Hence in any case

$$\ell_{\mathrm{J}}(\varphi_m(g)) \geq \tfrac{1}{2}(1 - \iota(\varphi_m(g))) \geq \tfrac{1}{4} - \varepsilon.$$

Furthermore Lemma 3.22 implies

$$\begin{aligned}
&\ell_{\mathrm{J}}(\varphi_m(g)\varphi_m(h)\varphi_m(gh)^{-1}) \\
&\quad = \ell_{\mathrm{J}}((\varphi(g)\varphi(h)\varphi(gh)^{-1})^{\otimes 2^m}) \\
&\quad \leq 2^m \ell_{\mathrm{J}}(\varphi(g)\varphi(h)\varphi(gh)^{-1}) \leq \varepsilon,
\end{aligned}$$

whenever $g, h \in E$.

Now suppose we are approximating with the rank length. To the pair (E,ε) choose m such that $f^m(1-\delta) \leq \frac{1}{2} + 4\varepsilon$, and $\varepsilon' < 2^{-m}\varepsilon$. Then there is an (E,ε')-homomorphism φ such that $\ell_{\mathrm{r}}(\varphi(g)) \geq \delta$ for all $g \in E$. We define φ_m as before, and additionally

$$\chi_k(g) := \varphi(g) \otimes \mathrm{id}_{n^{2^k}}, \qquad \psi_k(g) := \varphi_k(g) \oplus \chi_k(g).$$

If $\iota(\varphi(g)) \leq 1-\delta$, then we proceed as above to deduce $\ell_{\mathrm{r}}(\varphi_m(g)) \geq \ell_{\mathrm{J}}(\varphi_m(g)) \geq \frac{1}{4}-2\varepsilon$, and hence $\ell_{\mathrm{r}}(\psi_m(g)) \geq \frac{1}{8}-\varepsilon$. If $\iota(\varphi(g)) > 1-\delta$, then $\iota(\varphi(g)) = \iota_1(\varphi(g))$ is impossible, because this would imply $\ell_{\mathrm{r}}(\varphi(g)) \leq 1-\iota_1(\varphi(g)) < \delta$, contrary to the hypothesis. Thus we can assume $\iota(\varphi(g)) = \iota_\alpha(\varphi(g))$, where $\alpha \neq 1$. In this case $\iota_1(\varphi(g)) \leq 1-\iota_\alpha(\varphi(g)) < \delta$ and so $\ell_{\mathrm{r}}(\chi_m(g)) = \ell_{\mathrm{r}}(\varphi(g)) > \frac{1}{2}(1-\delta)$. At the same time $\ell_{\mathrm{r}}(\chi_m(g)) > \delta$, which implies $\ell_{\mathrm{r}}(\chi_m(g)) \geq \frac{1}{4}$. Thus $\ell_{\mathrm{r}}(\psi_m(g)) \geq \frac{1}{8}$ follows. Showing $\ell_{\mathrm{J}}(\psi_m(g)\psi_m(h)\psi_m(gh)^{-1}) \leq \varepsilon$ for $g,h \in E$ works as before.□

Note that Proposition 5.13 in [3], which is a special case of the previous theorem for matrices over $\mathbb{C}$ and the rank length, works with $\frac{1}{4}-\varepsilon$ instead of $\frac{1}{8}-\varepsilon$.

We will use the amplification properties proved in the previous theorem in § 6.

§ 5 LINEAR GROUP APPROXIMATION

Since we are in particular interested in finite matrix groups, we will use the abbreviation **(projectively)** q**-sofic** instead of (projectively) $\mathbb{F}_q$-sofic. If Γ can be embedded in an ultraproduct of groups $\mathrm{GL}_{n_i}(K_i)$ with respect to the Jordan length, where the K_i are finite fields of characteristic p_i and moreover $\lim_{\mathrm{u}} p_i = \infty$, then we call Γ **projectively** 0**-sofic**. If the rank length is used instead we will call Γ 0**-sofic**. We will also write **(projectively) prime sofic** or **(projectively) zero sofic** if Γ is (projectively) q-sofic, and q is a prime or 0, respectively. In the following we will examine the interplay of approximation with the rank length and Jordan length, and approximation in matrix groups over different fields.

The proof of the next proposition is clear.

PROPOSITION 3.24 *Let Γ be a (projectively) p-sofic group for infinitely many primes p. Then Γ is (projectively) zero sofic.*

As mentioned earlier there are still no non-sofic groups known. Since most groups we are dealing with obey even more general conditions than soficity, showing that an approximation property does not hold is under our general circumstances as hard as proving the existence of a non-sofic group. Hence unfortunately we will not be able to present counterexamples to the approximation properties of interest.

Showing the equivalence of different approximation properties is surely helpful, as already observed when we replaced the conjugacy length with the better behaved Jordan length. We would not voluntarily miss a chance to prove equivalences of this kind. Thus when in the following only one implication is proved, the other one is simply still unknown.

A metric ultraproduct of symmetric groups as in Theorem 3.10 is called a **universal sofic** group. We define **universal (projectively) linearly sofic** groups as ultraproducts of groups $\mathrm{GL}_{n_i}(K_i)$ equipped with the rank length (or Jordan length). Hence by Theorem 3.7 we deduce the following statement.

PROPOSITION 3.25 *Let Γ be a group. Then Γ is (projectively) linearly sofic if and only if it embeds into a universal (projectively) linearly sofic group.*

The whole terminology of group approximation with matrices we introduced so far can be used accordingly to define respective universal groups. (For an instance, a universal zero sofic group would be $\left(\prod \mathrm{GL}_{n_i}(K_i)\right)_{\mathrm{u}}$, where K_i is finite of characteristic p_i and $\lim_{\mathrm{u}} p_i = \infty$.) Proposition 3.25 can be reformulated for such a more restrictive setup.

PROPOSITION 3.26 *Let K be any field. Every sofic group is (projectively) K-sofic.*

P r o o f. The symmetric group S_n embeds into $\mathrm{GL}_n(K)$ as the subgroup of permutation matrices, independently of K. The generic length function on this subgroup inherited from the rank length or the Jordan length in $\mathrm{GL}_n(K)$ is asymptotically equivalent to the Hamming length by Proposition 2.10. □

Combining Proposition 3.24 and Proposition 3.26, we see that the class of sofic groups provides many examples of groups that are simultaneously prime sofic and zero sofic.

LEMMA 3.27 *Let Γ_1 and Γ_2 be groups, $E_i \subset \Gamma_i$ finite, and $\varphi\colon \Gamma_1 \to \mathrm{GL}_n(K)$ and $\psi\colon \Gamma_2 \to \mathrm{GL}_m(K)$ be $(E_i, \frac{1}{2}\varepsilon)$-homomorphisms with respect to the Jordan length for $i = 1, 2$. Then*

$$\zeta_{(g,h)} := \varphi_g \otimes \psi_h \in \mathrm{GL}_{nm}(K)$$

defines an $(E_1 \times E_2, \varepsilon)$-homomorphism on $\Gamma_1 \times \Gamma_2$.

Proof. Assume $\ell_J(\varphi_g) \geq \delta > 0$ and $\ell_J(\psi_h) \geq \delta$ for all $g \in E_1$, $h \in E_2$. Note that (3) in Lemma 3.22 is in the general form a very weak estimate, compared to (1). This is the reason why we are working with tensor products. By Lemma 3.22 for $g, g' \in E_1$ and $h, h' \in E_2$

$$\begin{aligned}\ell_J(\zeta_{(g,h)}\zeta_{(g',h')}\zeta^{-1}_{(g,h)(g',h')}) \\ &= \ell_J(\varphi_g\varphi_{g'}\varphi^{-1}_{gg'} \otimes \psi_h\psi_{h'}\psi^{-1}_{hh'}) \\ &\leq \ell_J(\varphi_g\varphi_{g'}\varphi^{-1}_{gg'}) + \ell_J(\psi_h\psi_{h'}\psi^{-1}_{hh'}) \\ &\leq \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon = \varepsilon.\end{aligned}$$

Assume one of $\iota(\varphi_g)$ and $\iota(\psi_h)$ is larger than $\frac{1}{2}$. Then we use Proposition 3.16 and Lemma 3.20 to estimate

$$\begin{aligned}\ell_J(\zeta_{(g,h)}) &= \ell_J(\varphi_g \otimes \psi_h) \\ &\geq \tfrac{1}{2}(1 - \iota(\varphi_g \otimes \psi_h)) \\ &\geq \tfrac{1}{2}(1 - \iota(\varphi_g)\iota(\psi_h) - (1-\iota(\varphi_g))(1-\iota(\psi_h))) \\ &= \tfrac{1}{2}\iota(\varphi_g)(1-\iota(\psi_h)) + \tfrac{1}{2}\iota(\psi_h)(1-\iota(\varphi_g)).\end{aligned}$$

Hence $\ell_J(\zeta_{(g,h)}) \geq \frac{1}{4}\ell_J(\varphi_g)$ or $\ell_J(\zeta_{(g,h)}) \geq \frac{1}{4}\ell_J(\psi_h)$, which is large enough. If $\iota(\varphi_g), \iota(\psi_h) \leq \frac{1}{2}$, then, also by Lemma 3.20,

$$\ell_J(\zeta_{(g,h)}) \geq \tfrac{1}{2}(1 - \iota(\varphi_g \otimes \psi_h)) \geq \tfrac{1}{4}. \qquad \square$$

Thus ζ is an $(E_1 \times E_2, \varepsilon)$-homomorphism.

THEOREM 3.28 *A group Γ is K-sofic if and only if it is projectively K-sofic.*

Proof. Let Γ be a K-sofic group and $\varphi\colon \Gamma \to \mathrm{GL}_n(K)$ an (E, ε)-homomorphism with respect to the rank length, where we may assume $\varepsilon < \frac{1}{2}$. Let

$$\psi_g := \varphi_g \oplus \mathrm{id}_n \in \mathrm{GL}_{2n}(K).$$

Then for every $g \in \Gamma$ evidently $\ell_r(\psi_g) \leq \frac{1}{2}$. By Corollary 2.25 $\ell_J(\psi_g) = \ell_r(\psi_g)$ follows. Hence $\ell_J(\psi_g) > \frac{\delta}{2}$ is true for all $g \in E$. Also, since $\varepsilon < \frac{1}{2}$, by Corollary 2.25 we have

$$\ell_J(\varphi_g\varphi_h\varphi^{-1}_{gh}) = \ell_r(\varphi_g\varphi_h\varphi^{-1}_{gh})$$

and consequently

$$\ell_{\mathrm{J}}(\psi_g \psi_h \psi_{gh}^{-1}) < \tfrac{1}{2}\varepsilon,$$

whenever $g, h \in E$. Thus Γ is projectively K-sofic.

Suppose conversely that Γ is projectively K-sofic. Let A^{T} denote the transpose of a matrix A, and $A^{-\mathrm{T}} := (A^{\mathrm{T}})^{-1}$. Then it is easily seen that $\ell_{\mathrm{r}}(A^{-\mathrm{T}}) = \ell_{\mathrm{r}}(A)$ and $\ell_{\mathrm{J}}(A^{-\mathrm{T}}) = \ell_{\mathrm{J}}(A)$. We choose an $(E, \frac{1}{2}\varepsilon)$-homomorphism φ, where $\varepsilon < \frac{1}{4}$ and define

$$\psi_g := \varphi_g \otimes \varphi_g^{-\mathrm{T}}.$$

If α is an eigenvalue of φ_g, then clearly α^{-1} is an eigenvalue of $\varphi_g^{-\mathrm{T}}$. Therefore $g \mapsto \varphi_g^{-\mathrm{T}}$ defines an $(E, \frac{1}{2}\varepsilon)$-homomorphism. Then by embedding Γ diagonally into $\Gamma \times \Gamma$ Lemma 3.27 shows that ψ is an (E, ε)-homomorphism with respect to the Jordan length. By the choice of ε, the prevalent eigenvalue of $\psi_g \psi_h \psi_{gh}^{-1}$ is 1 for all $g, h \in E$. Therefore we have

$$\ell_{\mathrm{r}}(\psi_g \psi_h \psi_{gh}^{-1}) = \ell_{\mathrm{J}}(\psi_g \psi_h \psi_{gh}^{-1}) \leq \varepsilon.$$

Moreover $\ell_{\mathrm{r}}(\psi_g) \geq \ell_{\mathrm{J}}(\psi_g)$ for $g \in E$. This shows that Γ is K-sofic. □

Theorem 3.28 will be of great use to reduce problems concerning projectively linearly sofic groups to linearly sofic groups. For example the proof of Theorem 3.23 could be done for linearly sofic groups and the statement for projectively sofic groups derived from Theorem 3.28. Of course this kind of reduction also works the other way round, but usually linearly sofic groups are a bit easier to handle.

The following proposition asserts that linear soficity is preserved when passing to larger fields. It is a good example of an application of Theorem 3.28. The converse statement is more complicated and will be explained afterwards.

PROPOSITION 3.29 *Let K be a field. If Γ is (projectively) K-sofic, then it is (projectively) L-sofic for any field L containing K.*

Proof. The claim for linearly sofic group follows directly from the definition of the rank length. The claim for projectively linearly sofic groups requires two applications of Theorem 3.28. □

LEMMA 3.30 *Let L/K be a finite field extension. Then matrices $A \in \mathrm{GL}_n(L)$ act as invertible linear transformations A' on $K^{n\cdot[L:K]}$ and*

$$\dim_K \ker(1-A') = [L:K] \cdot \dim_L \ker(1-A)$$

holds for all $A \in \mathrm{GL}_n(L)$.

Proof. We write $m := [L:K]$ and fix an isomorphism $\eta\colon L \to K^m$. We will denote the resulting isomorphism $L^n \cong K^{nm}$ also by η, and write A' for A acting on K^{nm}, i.e. $A'(x) = \eta(A(\eta^{-1}(x)))$ for $x \in K^{nm}$. This action is clearly linear and A' is invertible since η and A are bijective. Now $A'(x) = x$ if and only if $A(\eta^{-1}(x)) = \eta^{-1}(x)$ and $\eta(\langle y\rangle)$ is an m-dimensional subspace of K^{nm} for all $y \in L^n$ such that $A(y) = y$. Hence $\dim_K \ker(1-A') = m \dim_L \ker(1-A)$ follows as claimed. □

THEOREM 3.31 *Let L/K be an algebraic field extension and suppose Γ is (projectively) L-sofic. Then Γ is (projectively) K-sofic.*

Proof. Suppose Γ is L-sofic. We consider an (E,ε)-homomorphism $\varphi\colon \Gamma \to \mathrm{GL}_n(L)$. Let M be the set of all entries of matrices in $\varphi(E^2)$. Then $K(M)$ is a subfield of L and an algebraic extension of K of finite degree m. By Lemma 3.30 the elements of $\varphi(E^2)$ act on K^{nm} as linear transformations and their rank length remains unchanged. Thus we obtain an (E,ε)-homomorphism $\varphi\colon \Gamma \to \mathrm{GL}_{nm}(K)$, if we use this action for elements in $\varphi(E^2)$ and some arbitrary extension to $\Gamma\backslash E^2$. Hence Γ is K-sofic.

Theorem 3.28 yields the conclusion if Γ is projectively L-sofic. □

Let L/K be a field extension. Let $x = (x_1,\dots,x_m)$ be a vector in L^m. We say that $x' = (x'_1,\dots,x'_m)$ is a **specialization** of x over K if every polynomial $f \in K[X_1,\dots,X_m]$ vanishing at x does also vanish at x'. Moreover x' is an **algebraic** specialization if $K(x')/K$ is an algebraic extension.

THEOREM 3.32 *Let L/K be a field extension and suppose Γ is (projectively) L-sofic. Then Γ is (projectively) K-sofic.*

Proof. We consider an (E,ε)-homomorphism $\varphi\colon\Gamma\to\mathrm{GL}_n(L)$. Let X be the set of all entries of matrices in $\varphi(E^2)$. For every $g\in E$ let a_g be a maximal $k_g\times k_g$-submatrix of $1-\varphi_g$ such that $\det(a_g)\neq 0$, i.e. $k_g=\mathrm{rk}(1-\varphi_g)$. Then there is $\alpha_g\in L$ such that $\alpha_g\det(a_g)-1=0$. Let Y be the set of all α_g for $g\in E$. We order the elements of $X\cup Y$ as a vector x. Then by Theorem 7 in [29], Chapter II, there exists an algebraic specialization x' of x. This implies that $K(x')$ is an algebraic extension of K. We write ψ_g for the matrix in $K(x')$ obtained by replacing elements in x with appropriate elements in x', and a'_g for the submatrix of $1-\psi_g$ corresponding to a_g. Then every submatrix of $1-\psi_g$ larger than a'_g has zero determinant, and $\alpha'_g\det(a'_g)-1=0$, since the determinant is a polynomial in matrix entries. Therefore $\det(a'_g)\neq 0$, $\mathrm{rk}(1-\psi_g)=k_g$ and consequently $\ell_{\mathrm{r}}(\psi_g)=\ell_{\mathrm{r}}(\varphi_g)$. By the same reasoning $\ell_{\mathrm{r}}(\psi_g\psi_h\psi_{gh}^{-1})\leq\varepsilon$. If we define ψ_g arbitrary for $g\notin E^2$, then ψ is an (E,ε)-homomorphism into $\mathrm{GL}_n(K(x'))$. We have thus shown that Γ is $K(x')$-sofic, and since $K(x')/K$ is algebraic, by Theorem 3.31 Γ is K-sofic.

Now Theorem 3.28 proves the claim for projectively linearly sofic groups. □

COROLLARY 3.33 *Let $q=p^k$, where p is a prime power. Then every q-sofic group is p-sofic.*

COROLLARY 3.34 *Let Γ have the approximation property in $\mathcal{GL}_{\mathrm{fin}}$ with the rank length. Then Γ is p-sofic for p a prime or 0.*

Proof. We can embed Γ into an ultraproduct of groups $\mathrm{GL}_{n_i}(K_i)$, where the K_i are finite fields. If K_i has characteristic p_i and $\lim_{\mathrm{u}} p_i=\infty$, then Γ is by definition 0-sofic. If otherwise $p_i\leq C\ [\mathrm{u}]$ for some constant C, then by Proposition 3.1 $p_i=p\ [\mathrm{u}]$. Proposition 3.3 shows that $\left(\prod\mathrm{GL}_{n_i}(K_i)\right)_{\mathrm{u}}$ is isomorphic to $\left(\prod_{i\colon \mathrm{char}(K_i)=p}\mathrm{GL}_{n_i}(K_i)\right)_{\mathrm{u}}$. By Theorem 3.31 we can replace all K_i of characteristic p with the field $\mathbb{F}_p$ to show that Γ is p-sofic. □

In [3], Theorem 8.2 it was proved that a $\mathbb{C}$-sofic group is prime sofic or zero sofic. We adopt exactly the same argument for arbitrary fields:

THEOREM 3.35 *Let Γ be (projectively) linearly sofic. Then Γ has the $\mathcal{GL}_{\mathrm{fin}}$-approximation property with respect to the rank length (or Jordan length, respectively.)*

Proof. We first consider approximation with the rank length. Let $\varepsilon > 0$ and $E \subset \Gamma$ be a finite subset. Then we find an (E,ε)-homomorphism $\varphi\colon \Gamma \to \mathrm{GL}_n(K)$ for appropriate n and a field K. We can assume without loss of generality that $\varphi_g = 1$ for all $g \notin E^2$. For every $g \in E$ let a_g be a maximal submatrix of $1-\varphi_g$ such that $\delta_g := \det(a_g) \neq 0$. That is a_g is a $k \times k$-matrix, where $\ell_{\mathrm{r}}(\varphi_g) = \frac{k}{n}$.

Let R be the unital subring of K generated by all entries of matrices in $\varphi(E)$. Then R is a finitely generated algebra over the prime field of K if $\mathrm{char}(K) > 0$ or a finitely generated $\mathbb{Z}$-algebra otherwise. Since $\mathbb{F}_p$ and $\mathbb{Z}$ are Jacobson rings, by Theorem 4.19 in [13] R is such. It is clear that φ takes values in $\mathrm{GL}_n(R)$. Since R is a Jacobson ring, there is a maximal ideal m in R such that

$$\left\{\delta_g \,\middle|\, g \in E\right\} \cap m = \varnothing.$$

Since m is a maximal ideal, R/m is a field, and because R/m is finitely generated as a ring, it is a finite field. Now we let $\psi := \pi \circ \varphi$, where π denotes application of the quotient morphism $R \to R/m$ to the entries of matrices in $\mathrm{GL}_n(R)$. Then for arbitrary $g,h \in E$

$$\ell_{\mathrm{r}}(\psi_g\psi_h\psi_{gh}^{-1}) = \frac{\mathrm{rk}(\psi_{gh} - \psi_g\psi_h)}{n} \leq \frac{\mathrm{rk}(\varphi_{gh} - \varphi_g\varphi_h)}{n} = \ell_{\mathrm{r}}(\varphi_g\varphi_h\varphi_{gh}^{-1}).$$

Since the choice of m took the determinants δ_g into account, for every $g \in E$

$$\ell_{\mathrm{r}}(\psi_g) = \frac{\mathrm{rk}(1-\psi_g)}{n} = \frac{\mathrm{rk}(1-\varphi_g)}{n} = \ell_{\mathrm{r}}(\varphi_g).$$

Of course $\psi_1 = 1$. Hence $\psi\colon \Gamma \to \mathrm{GL}_n(R/m)$ is an (E,ε)-homomorphism.

Theorem 3.28 proves the claim for projectively linearly sofic groups. □

COROLLARY 3.36 *Every (projectively) linearly sofic group is (projectively) p-sofic, where p is a prime or $p = 0$.*

We are now confronted with the two similar Theorems 3.32 and 3.35. The latter is somewhat stronger insofar as it reduces approximation to finite fields, but it lacks the virtue of the former of preserving the field characteristic.

We are left with some open questions. The most important one in this context surely is whether linearly sofic groups are sofic. It is also unclear if p-sofic groups are q-sofic for different primes p and q, or if prime sofic or zero sofic groups are $\mathbb{Q}$-sofic. These questions were also adressed in [3], Question 8.6. A positive answer combined with Theorem 7.4, ibid. would for example solve Kaplansky's Direct Finiteness Conjecture for linearly sofic groups (see also ibid, Question 7.9.)

§ 6 THE CLASS OF LINEARLY SOFIC GROUPS

In [16] it was proved that the class of sofic groups is closed under taking subgroups, direct products, direct limits, inverse limits, free products and extensions by amenable groups. Later Collins and Dykema proved in [11] that free products of sofic groups amalgamated over monotileably amenable groups are sofic. The unpleasant restriction of monotileability was removed by Păunescu in [35] and Elek and Szabó in [17]. We shall show that similar conclusions as in [16] are true for linearly sofic groups, but we will not go as far as treating amalgamated products.

PROPOSITION 3.37 *Let K be a field. The class of K-sofic groups is closed under taking subgroups, inverse limits, direct products and direct limits.*

Proof. The claim concerning subgroups and inverse limits is an immediate consequence of Proposition 3.13.

Now suppose Γ is a direct limit of K-sofic groups Γ_i. By Proposition 3.15 and Theorem 3.23 Γ is K-sofic.

By Lemma 3.22 $\mathrm{GL}_n(K) \times \mathrm{GL}_m(K)$ embeds isometrically into $\mathrm{GL}_{n+m}(K)$ when using the mapping $(A,B) \mapsto A \oplus B$. Thus, with an appeal to induction, Proposition 3.14 shows that direct products of K-sofic groups are K-sofic. □

LEMMA 3.38 *Let K be a field and $\pi \in S_n$ a permutation without fixed points. Let $A_i \in \mathrm{GL}_m(K)$ for $i = 1 \dots n$, and P_π be the permutation matrix in $\mathrm{GL}_n(K)$ corresponding to π. We define $A \in \mathrm{GL}_{nm}(K)$ by $A(e_i \otimes e_j) := e_i \otimes A_i(e_j)$ and linear extension. Then $\mathrm{rk}(1-(P_\pi \otimes \mathrm{id}_m) \circ A) \geq \frac{1}{2}nm$.*

Proof. The matrix $P_\pi \otimes \mathrm{id}_m$ is blockdiagonal, where every block corresponds to a cycle of π. Hence it suffices to assume that π consists of a single cycle. Then it is an elementary observation that if n is even, $1-(P_\pi \otimes \mathrm{id}_m)\circ A$ has a submatrix $\mathrm{id}_{\frac{1}{2}nm}$. If n is odd there is a $\frac{1}{2}(n+1)m \times \frac{1}{2}(n+1)m$-submatrix which is a block matrix having blocks id_m on the diagonal and one block $-A_j$ for some j. In both cases the determinant of the submatrix is non-zero and the claim follows.□

The proof of the following proposition is a linearized version of the proof of Item 3 in Theorem 1, [16]. A variant of the statement appears in [3], Theorem 9.3 for $\mathbb{C}$-sofic groups.

PROPOSITION 3.39 *Let K be a field and Γ a group such that $N \lhd \Gamma$ is K-sofic and Γ/N is amenable. Then Γ is K-sofic.*

Proof. Let E be a finite subset of Γ and $\varepsilon > 0$. If $g \in \Gamma$, we write $\overline{g} := gN \in \Gamma/N$. Let $\sigma : \Gamma/N \to \Gamma$ be a section of the canonical projection, i.e. $\sigma(\overline{g})^{-1}g \in N$ for all $g \in \Gamma$. We can choose a set $A \subset \Gamma$ with the property $\sigma(\overline{A}) = A$ and $|\overline{g}\overline{A}\Delta\overline{A}| \leq \frac{1}{3}\varepsilon|\overline{A}|$ for all $g \in E$, i.e. $\overline{A}$ is an $(\overline{E}, \frac{1}{3}\varepsilon)$-Følner set. Let $D := N \cap (A^{-1}EA)$ and $\psi : N \to \mathrm{GL}_n(K)$ be a $(D, \frac{1}{3}\varepsilon)$-homomorphism. We denote the standard basis vectors of K^n and K^A by e_i, $i = 1\ldots n$ and e_a, $a \in A$, respectively. If $\overline{ga} \in \overline{A}$, then $\sigma(\overline{ga}) \in A$. Hence we can define the permutation $\pi_g \in S_A$ by

$$\pi_g(a) := \sigma(\overline{ga})$$

for $\overline{ga} \in \overline{A}$ and by taking arbitrary values for the remaining elements. Because $\sigma(\overline{ga})^{-1}ga \in N$ for all $a \in A$ we can also define $\varphi_g \in \mathrm{GL}(K^n \otimes K^A)$ by

$$\varphi_g(e_1 \otimes e_a) := \psi_{\sigma(\overline{ga})^{-1}ga}(e_i) \otimes e_{\pi_g(a)}$$

and linear extension. Then we extend φ arbitrarily to a mapping $\Gamma \to \mathrm{GL}(K^n \otimes K^A)$. We shall show that φ is in fact an (E, ε)-homomorphism.

Because σ is a section, $\psi_{\sigma(\overline{a})^{-1}a} = \psi_{a^{-1}a} = \psi_1 = 1$. Hence for all $i = 1\ldots n$ and $a \in A$

$$\varphi_1(e_i \otimes e_a) = \psi_{\sigma(\overline{a})^{-1}a}(e_i) \otimes e_{\sigma(\overline{a})} = e_i \otimes e_a.$$

Now assume $g \in E \setminus \{1\}$ and $g \notin N$. This implies $\overline{g} \neq 1$ and $a \neq \sigma(\overline{ga})$ for all $a \in A$. If $\overline{ga} \in \overline{A}$, then $\pi_g(a) = \sigma(\overline{ga})$. The number of elements a such that

$\overline{ga} \in \overline{A}$ is more than $(1-\frac{1}{3}\varepsilon)|A|$. Hence the number of fixed points of π_g is less than $\frac{1}{3}\varepsilon|A|$. By Lemma 3.38

$$\mathrm{rk}(1-\varphi_g) \geq \tfrac{1}{2}n \cdot (1-\tfrac{1}{3}\varepsilon)|A|,$$

whence $\ell_{\mathrm{r}}(\varphi_g) \geq \frac{1}{4}$. If otherwise $g \in N \cap (E \setminus \{1\})$, then $\overline{g} = 1$ and $\varphi_g(e_i \otimes e_a) = \psi_{a^{-1}ga}(e_i) \otimes e_a$. Now $\mathrm{rk}(1-\varphi_g) = \sum_{a\in A} \mathrm{rk}(1-\psi_{a^{-1}ga})$ and consequently

$$\ell_{\mathrm{r}}(\varphi_g) = \frac{\sum_{a\in A} \ell_{\mathrm{r}}(\psi_{a^{-1}ga})}{|A|} \geq \delta.$$

For the remaining estimate let $g, h \in E$ and assume $\overline{ha}, \overline{g}\overline{ha} \in \overline{A}$ for some $a \in A$. This is not possible for at most $\frac{2}{3}\varepsilon|A|$ elements $a \in A$. Since $\sigma(\overline{g}\sigma(\overline{ha})) = \sigma(\overline{g}\overline{ha})$,

$$\begin{aligned} \varphi_g\varphi_h(e_i \otimes e_a) &= \varphi_g(\psi_{\sigma(\overline{ha})^{-1}ha}(e_i) \otimes e_{\sigma(\overline{ha})}) \\ &= \psi_{\sigma(\overline{g}\overline{ha})^{-1}g\cdot\sigma(\overline{ha})}\psi_{\sigma(\overline{ha})^{-1}ha}(e_i) \otimes e_{\sigma(\overline{g}\overline{ha})}. \end{aligned}$$

Furthermore

$$\varphi_{gh}(e_i \otimes e_a) = \psi_{\sigma(\overline{g}\overline{ha})^{-1}gha}(e_i) \otimes e_{\sigma(\overline{g}\overline{ha})}.$$

In particular $\pi_g\pi_h\pi_{gh}^{-1}$ has at least $(1-\frac{2}{3}\varepsilon)|A|$ fixed points. Because $\overline{ha}$ and $\overline{g}\overline{ha}$ were assumed to be in $\overline{A}$, the elements $\sigma(\overline{ha})^{-1}ha$ and $\sigma(\overline{g}\overline{ha})^{-1}g \cdot \sigma(\overline{ha})$ are in D. Now we can exploit that ψ is an almost homomorphism and obtain

$$\begin{aligned} \mathrm{rk}(\varphi_g\varphi_h - \varphi_{gh}) &\leq \sum_{\overline{ha}, \overline{g}\overline{ha} \in \overline{A}} \mathrm{rk}(\psi_{\sigma(\overline{g}\overline{ha})^{-1}g\cdot\sigma(\overline{ha})}\psi_{\sigma(\overline{ha})^{-1}ha} - \psi_{\sigma(\overline{g}\overline{ha})^{-1}gha}) + \tfrac{2}{3}\varepsilon|A| \cdot n \\ &\leq |A| \cdot \tfrac{1}{3}\varepsilon n + \tfrac{2}{3}\varepsilon|A| \cdot n, \end{aligned}$$

and consequently

$$\ell_{\mathrm{r}}(\varphi_g\varphi_h\varphi_{gh}^{-1}) \leq \tfrac{1}{3}\varepsilon + \tfrac{2}{3}\varepsilon = \varepsilon.$$

The proof that φ is an (E,ε)-homomorphism is now complete. □

LEMMA 3.40 *Let Γ be a group with the $\mathcal{G}$-approximation property. Then for any finite subset $E \subset \Gamma$ and $\varepsilon > 0$ there is an (E,ε)-homomorphism $\varphi\colon \Gamma \to G \in \mathcal{G}$ such that $\varphi(1) = 1$ and $\varphi(g^{-1}) = \varphi(g)^{-1}$ for all $g \in E$ not of order 2.*

Proof. Let $F := (E \cup E^{-1} \cup \{1\})^2$ and $\psi\colon \Gamma \to G$ be an $(F, \frac{1}{2}\varepsilon)$-homomorphism into a group $G \in \mathcal{G}$ with invariant pseudo length function ℓ. By Proposition 3.9 without loss of generality $\psi(1) = 1$. We partition F into three subsets as follows: Let F_0 be the set of all $g \in F$ of order 2 or $g = 1$. Then we partition $F \setminus F_0$ into F_1 and F_{-1} such that $g \in F_{-1}$ implies $g^{-1} \in F_1$, or equivalently $g \in F_1$ implies $g^{-1} \in F_{-1}$. We define

$$\varphi(g) := \begin{cases} \psi(g), & g \in F_0 \cup F_1, \\ \psi(g^{-1})^{-1}, & g \in F_{-1}. \end{cases}$$

Then obviously $g \in F_{-1}$ implies $\varphi(g)^{-1} = \varphi(g^{-1})$, and the same holds for $g \in F_1$, since in this case $g^{-1} \in F_{-1}$.

We must prove that φ is an (E, ε)-homomorphism. It is enough to show that $g, h \in E$ implies $\ell(\varphi(g)\varphi(h)\varphi(gh)^{-1}) \le \varepsilon$. The case of $g, h, gh \in F_0 \cup F_1$ is clear. The case of $g, h, gh \in F_{-1}$ reduces to

$$\ell(\varphi(g)\varphi(h)\varphi(gh)^{-1}) = \ell(\psi(h^{-1})\psi(g^{-1})\psi(h^{-1}g^{-1})^{-1}) \le \tfrac{1}{2}\varepsilon,$$

where we used the invariance of ℓ. Let $g, h \in F_0 \cup F_1$, $gh \in F_{-1}$. Then

$$\begin{aligned} &\ell(\varphi(g)\varphi(h)\varphi(gh)^{-1}) \\ &\quad = \ell(\psi(g)\psi(h) \cdot \psi(gh)^{-1}\psi(gh) \cdot \psi((gh)^{-1})) \\ &\quad \le \ell(\psi(g)\psi(h)\psi(gh)^{-1}) + \ell(\psi(gh)\psi((gh)^{-1})) \\ &\quad \le \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon = \varepsilon. \end{aligned}$$

If $g \in F_0 \cup F_1$ and $h, gh \in F_{-1}$ we estimate

$$\begin{aligned} &\ell(\varphi(g)\varphi(h)\varphi(gh)^{-1}) \\ &\quad = \ell(\psi(g) \cdot \psi(g^{-1})\psi(g^{-1})^{-1} \cdot \psi(h^{-1})^{-1}\psi((gh^{-1})) \\ &\quad \le \ell(\psi(g)\psi(g^{-1})) + \ell(\psi(h^{-1})\psi(g^{-1})\psi(h^{-1}g^{-1})^{-1}) \\ &\quad \le \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon = \varepsilon. \end{aligned}$$

Finally assume $g, gh \in F_0 \cup F_1$ and $h \in F_{-1}$. Then

$$\begin{aligned}
&\ell(\varphi(g)\varphi(h)\varphi(gh)^{-1})\\
&\qquad = \ell(\psi(g)\psi(h^{-1})^{-1} \cdot \psi(h)^{-1}\psi(h) \cdot \psi(gh)^{-1})\\
&\qquad \leq \ell(\psi(h^{-1})\psi(h)) + \ell(\psi(g)\psi(h)\psi(gh)^{-1})\\
&\qquad \leq \tfrac{1}{2}\varepsilon + \tfrac{1}{2}\varepsilon = \varepsilon
\end{aligned}$$

holds. The remaining cases follow analogously. □

LEMMA 3.41 *Let K be a field and Γ_1 and Γ_2 finitely generated K-sofic groups. Then the free product $\Gamma_1 * \Gamma_2$ is K-sofic.*

Proof. Suppose Γ_1 and Γ_2 are generated by finite symmetric sets A and B, respectively. Let $\varphi\colon \Gamma_1 \to \mathrm{GL}_n(K)$ be an $(A^r, \frac{1}{2}\varepsilon)$-homomorphism and $\psi\colon \Gamma_2 \to \mathrm{GL}_m(K)$ a $(B^r, \frac{1}{2}\varepsilon)$-homomorphism. By Lemma 3.40 we can assume without loss of generality that $\varphi_1 = 1$, $\psi_1 = 1$, and $\varphi_{g^{-1}} = \varphi_g^{-1}$ and $\psi_{h^{-1}} = \psi_h^{-1}$ for all $g \in A^{2r}$ and $h \in B^{2r}$ not of order 2. We define a mapping $\varphi'\colon \Gamma_1 \to \mathrm{GL}_{2n}(K)$ by

$$\varphi'_g := \begin{pmatrix} \varphi_g & \\ & \varphi_g \end{pmatrix}$$

if $g^2 \neq 1$ and

$$\varphi'_g := \begin{pmatrix} & \varphi_g^{-1} \\ \varphi_g & \end{pmatrix}$$

otherwise. We define $\psi'\colon \Gamma_2 \to \mathrm{GL}_{2m}(K)$ analogously. Note that φ' and ψ' no longer need to be almost homomorphisms. Nevertheless $\varphi'_{g^{-1}} = (\varphi'_g)^{-1}$ holds for all $g \in A^{2r}$ and $\psi'_{h^{-1}} = (\psi'_h)^{-1}$ holds for all $h \in B^{2r}$.

Now consider the subgroup of $\mathrm{GL}_{2n}(K)$ generated by $\left\{\varphi'_g \,\middle|\, g \in A^{2r}\right\}$. By Malcev's Theorem this group is residually finite and hence there is a finite group H_1 and a homomorphism

$$\pi_1\colon \left\langle \left\{\varphi'_g \,\middle|\, g \in A^{2r}\right\}\right\rangle \to H_1,$$

the restriction of which to $\left\{\varphi'_g \,\middle|\, g \in A^{2r}\right\}$ is injective. Analogously there are H_2 and

$$\pi_2\colon \left\langle \left\{\psi'_h \,\middle|\, h \in B^{2r}\right\}\right\rangle \to H_2.$$

Since the free product of finite groups is residually finite by [25], Theorem 4.1, there is a finite group G and a homomorphism $\pi\colon H_1 * H_2 \to G$ such that $\pi(h_1 g_1 \dots h_k g_k) \neq 1$ if $h_1 g_1 \dots h_k g_k$ is a reduced word in $H_1 * H_2$, $h_i \in H_1$ and $g_i \in H_2$, and $k \leq 2r$. For elements $g \in A^{2r}$ or $h \in B^{2r}$ we write

$$\overline{g} := \pi(\pi_1(\varphi'_g)) \in G, \qquad \overline{h} := \pi(\pi_2(\psi'_h)) \in G.$$

To summarize, $a, a_i \in A^{2r}$ and $b, b_i \in B^{2r}$ imply $\overline{a^{-1}} = \overline{a}^{-1}$ and $\overline{b^{-1}} = \overline{b}^{-1}$, and $\overline{a_1}\,\overline{b_1} \dots \overline{a_k}\,\overline{b_k} \neq 1$, whenever $k \leq 2r$ and $a_i \neq 1$, $b_j \neq 1$ for $i \neq 1$, $j \neq k$.

Consider the vector space

$$V := K^G \otimes (K^n \oplus K^m)$$

with the basis of standard vectors $e_x \otimes e_i$, where $x \in G$, $i = 1 \dots n+m$. If $g \in A^{2r}$ we define

$$\tilde{\varphi}_g := \mathrm{id}_{K^G} \otimes (\varphi_g \oplus \mathrm{id}_m)$$

and similarly for $h \in B^{2r}$

$$\tilde{\psi}_h := \mathrm{id}_{K^G} \otimes (\mathrm{id}_n \oplus \psi_h).$$

Let $g = a_1 b_1 \dots a_k b_k$ be a reduced word in $(A \cup B)^{2r} \setminus \{1\} \subset \Gamma_1 * \Gamma_2$, where $a_i \in \Gamma_1$ and $b_i \in \Gamma_2$ for all $i = 1 \dots k$. Then $a_1 \dots a_k \in A^{2r}$ and $b_1 \dots b_k \in B^{2r}$. We let

$$\sigma_g(e_x \otimes e_i) := e_{\overline{a_1}\,\overline{b_1} \dots \overline{a_k}\,\overline{b_k} x} \otimes e_i.$$

Then σ_g commutes with $\tilde{\varphi}_a$ and $\tilde{\psi}_b$ for all $g \in (A \cup B)^{2r}$, $a \in A^{2r}$ and $b \in B^{2r}$. Now we define ζ_g by

$$\zeta_g := \sigma_g \circ \tilde{\varphi}_{a_1 \dots a_k} \circ \tilde{\psi}_{b_1 \dots b_k}$$

and linear extension. We let $\zeta_1 := \mathrm{id}_V$ and extend $\zeta\colon g \mapsto \zeta_g$ arbitrarily to the whole of Γ to obtain a mapping $\zeta\colon \Gamma \to \mathrm{GL}(V)$.

Let $g = a_1 b_1 \dots a_k b_k$ and $h = c_1 d_1 \dots c_l d_l$ be reduced words in $(A \cup B)^r \setminus \{1\} \subset \Gamma_1 * \Gamma_2$. We abbreviate $a := a_1 \dots a_k$, $b := b_1 \dots b_k$, $c := c_1 \dots c_l$ and $d := d_1 \dots d_l$. Suppose when multiplying g and h cancellations occur, i.e.

$$gh = a_1 b_1 \dots b_{k-s-1}(a_{k-s} c_{1+s}) d_{1+s} \dots c_l d_l$$

or

$$gh = a_1 b_1 \dots a_{k-t}(b_{k-t} d_{1+t}) c_{2+t} \dots c_l d_l.$$

Then without loss of generality in the first case $b_k = 1$, $a_{k-i} = c_{1+i}^{-1}$ and $b_{k-j} = d_j^{-1}$ for all $i = 0 \dots s$ and $j = 1 \dots s$. Therefore $\overline{b_k} = 1$, $\overline{a_{k-i}} = \overline{c_{1+i}}^{-1}$ and $\overline{b_{k-j}} = \overline{d_j}^{-1}$ in G. Thus when multiplying $\overline{a_1}\,\overline{b_1} \dots \overline{a_k}\,\overline{b_k}$ and $\overline{c_1}\,\overline{d_1} \dots \overline{c_l}\,\overline{d_l}$, the same cancellations as in gh occur (and maybe more). This means $\sigma_g \sigma_h = \sigma_{gh}$.

Now by the definition of ζ we readily obtain

$$\zeta_g \zeta_h - \zeta_{gh} = \sigma_{gh} \circ (\tilde{\varphi}_a \circ \tilde{\psi}_b \circ \tilde{\varphi}_c \circ \tilde{\psi}_d - \tilde{\varphi}_{ac} \circ \tilde{\psi}_{bd}).$$

Because σ_{gh} has full rank,

$$\begin{aligned}
\ell_{\mathrm{J}}(\zeta_g \zeta_h - \zeta_{gh}) &= \ell_{\mathrm{J}}(\tilde{\varphi}_a \circ \tilde{\psi}_b \circ \tilde{\varphi}_c \circ \tilde{\psi}_d - \tilde{\varphi}_{ac} \circ \tilde{\psi}_{bd}) \\
&= \ell_{\mathrm{J}}(\varphi_a \oplus \psi_b \circ \varphi_c \oplus \psi_d - \varphi_{ac} \oplus \psi_{bd}) \\
&\leq \frac{n}{n+m} \ell_{\mathrm{r}}(\varphi_a \varphi_c - \varphi_{ac}) + \frac{m}{n+m} \ell_{\mathrm{r}}(\psi_b \psi_d - \psi_{bd}),
\end{aligned}$$

where the inequality follows from (1) in Lemma 3.22. Since $a, c \in A^r$ and $b, d \in B^r$, the right hand side is less than ε.

It is clear that σ_g acts as a permutation matrix modulo $K^n \oplus K^m$. If $1 \neq g = a_1 b_1 \dots a_k b_k$ is a reduced word of length not more than $2r$ in letters from $A \cup B$, by construction $\overline{a_1}\,\overline{b_1} \dots \overline{a_k}\,\overline{b_k} \neq 1$ in G. Since the action of σ_g is determined by the permutation action of $\overline{a_1}\,\overline{b_1} \dots \overline{a_k}\,\overline{b_k}$ on G, we can use Lemma 3.38 to conclude that $\ell_{\mathrm{J}}(\zeta_g) \geq \frac{1}{2}$. □

Theorem 3.42 *Let K be a field. Then the free product of K-sofic groups is K-sofic.*

Proof. Let Γ_1 and Γ_2 be K-sofic groups. Then $\Gamma_i * \Gamma_2$ is a direct limit of groups $\Gamma_1^{(i)} * \Gamma_2^{(i)}$, where $\Gamma_j^{(i)}$ is a finitely generated subgroup of Γ_j for every i and $j = 1, 2$. As subgroups of Γ_j, the groups $\Gamma_j^{(i)}$ are K-sofic, and since they are finitely generated, by Lemma 3.41 $\Gamma_1^{(i)} * \Gamma_2^{(i)}$ is K-sofic for all i. At last Proposition 3.37 shows that the direct limit $\Gamma_1 * \Gamma_2$ is K-sofic. □

4 Continuous regular rings and linear group approximation

§ 1 Continuous geometries and continuous regular rings

The theory of continuous geometries was developed by von Neumann in the late thirties of the 20th century as an attempt at a useful formulation of quantum mechanics. It includes a more algebraic counterpart to the rings of operators we nowadays call von Neumann algebras. Von Neumann's work was later edited by Halperin into the text book [42]. We shall also cite from Maeda's more recent book [32] which presents a more systematic approach. The following introduction serves to build acquaintance with the most basic definitions and fundamental results.

Let L be a **lattice**, i.e. a non-empty set with commutative and associate binary operations **join** and **meet**, written $\vee$ and $\wedge$, respectively, such that $(a \vee b) \wedge a = a$ and $(a \wedge b) \vee a = a$ hold for all $a, b \in L$. Then L is called **modular** if it satisfies the **modular law**

$$(a \wedge c) \vee (b \wedge c) = ((a \wedge c) \vee b) \wedge c$$

for all $a, b, c \in L$. The operations $\vee$ and $\wedge$ induce an ordering $\leq$ on L. In this ordering $a \wedge b$ is the greatest lower bound of a and b, and $a \vee b$ is the least upper bound. If L contains a least upper bound to all its elements, we write 1 for this element. A greatest lower bound to all elements of L is denoted by $\mathtt{0}$. If $\mathtt{0}$ and 1 exist in L we call b a **complement** of a if $a \wedge b = \mathtt{0}$ and $a \vee b = 1$. A lattice where every element has a complement is a **complemented lattice**. A lattice is called **irreducible** if it is not isomorphic to a direct product of lattices containing at least two elements.

A lattice L is **complete** if the binary operations $\vee$ and $\wedge$ extend to infinitary operations compatible with the ordering of L. That is for all systems $X = \{x_i \mid \forall i \in I : x_i \in L\}$ in L, where I is an index set

$$\bigwedge_{i \in I} x_i, \qquad \bigvee_{i \in I} x_i$$

exist, and $\bigwedge_{i \in I} x_i$ is the greatest lower bound of X while $\bigvee_{i \in I} x_i$ is the least upper bound of X. Note that a complete lattice automatically contains 0 and 1. If I is a directed index set consider increasing systems $\{x_i \mid i \in I\}$, i.e. $i \leq j$ implies $x_i \leq x_j$, and decreasing systems satisfying $x_j \leq x_i$ for $i \leq j$. We call a complete lattice L **continuous** if for all $x \in L$, increasing systems $\{x_i \mid i \in I\}$ and decreasing systems $\{y_i \mid i \in I\}$ the infinitary distributive laws

$$x \wedge \bigvee_{i \in I} x_i = \bigvee_{i \in I}(x \wedge x_i), \qquad x \vee \bigwedge_{i \in I} y_i = \bigwedge_{i \in I}(x \vee y_i)$$

hold.

A continuous complemented modular lattice with at least two elements is called a **continuous geometry**.

Let R be a unital ring. Then R is called **von Neumann regular** if for every $a \in R$ there exists a **pseudo inverse** $x \in R$, i.e. $a = axa$ holds. We use here the shorter term **regular** instead, although this collides with notation in other areas of ring theory. The typical example of regular rings are matrix rings and in fact these are the basis of our considerations. It can be shown that the set $L(R)$ of all principal right ideals of a regular ring R, where the meet operation is intersection and the join operation is the sum of ideals, is a complemented modular lattice. (Confer [32], Chapter VI.) If $L(R)$ is a continuous lattice, then R is called a **continuous ring**.

Now we take a radical shortcut from the first definitions right to the main result, the so called *Coordinization Theorem*. Its deduction takes most of the space in [42] and [32]. We have to explain the following technical terms first. Two elements a and b in a complemented lattice L with 0 and 1 are called **perspective** if they have a common complement, i.e. there exists $c \in L$ such that $a \vee c = b \vee c = 1$ and $a \wedge c = b \wedge c = 0$. If L contains $a_1, \ldots, a_n$ such that $a_1 \vee \ldots \vee a_n = 1$ and a_i is perspective to a_j for $i, j = 1 \ldots n$, then L has **order** n. Note that the order of L is not unique. A regular ring R is said to have **order** n if $L(R)$ has order n.

> **THEOREM 4.1 (von Neumann's Coordinization Theorem)** *If L is a continuous geometry of order $n \geq 4$, then there exists up to isomorphism a unique continuous regular ring R such that L and $L(R)$ are isomorphic as continuous geometries.*

We will only have to deal with irreducible continuous geometries and, as the previous theorem suggests, irreducible rings. (A ring R is called **irreducible** if it is not the direct sum of two rings not isomorphic to $\{0\}$ or R.)

PROPOSITION 4.2 *Let R be a continuous regular ring. Then R is irreducible if and only if it is simple, if and only if the center of R is a field.*

Proof. Confer [32], Chapter VII, Hilfssatz 3.1 and Chapter IV, Satz 3.6. □

An important tool in the study of regular rings is provided by the following theorem, which is a direct corollary of [32], Chapter VII, Satz 1.2.

THEOREM 4.3 *Let R be an irreducible continuous regular ring. There exists a rank function $\rho\colon R \to [0,1]$ such that $ef = fe = 0$ implies $\rho(e+f) = \rho(f+e)$ for all idempotent elements $e, f \in R$.*

Note that in the context of regular rings the additional property concerning idempotents is included in the definition of a rank function. We will use this definition in this section.

The next important results about rank functions apart from mere existence are the following.

THEOREM 4.4 ([32], Chapter VII, Satz 5.3) *Let R be a regular ring with a rank function ρ. If R is a complete metric space in the metric given by $d(a,b) = \rho(a-b)$, then R is a continuous regular ring.*

THEOREM 4.5 ([32], Chapter VII, Satz 2.2) *An irreducible regular ring R is a continuous regular ring if and only if it possesses a rank function ρ and is complete as a metric space in the metric given by $d(a,b) = \rho(a-b)$. In this case ρ is unique.*

We conclude the introduction by presenting the continuous regular rings of main interest to us.

THEOREM 4.6 *Let $M_{n_i}(K_i)$ be the algebra of all $n_i \times n_i$-matrices over a field K_i for all i in the index set I. Consider $M_{n_i}(K_i)$ as a metric space, where the metric is induced by the normalized rank. Then the ring*

$$(\boldsymbol{R})_{\mathrm{u}} := \left(\prod M_{n_i}(K_i)\right)_{\mathrm{u}}$$

is a simple continuous regular ring with unique rank function ρ given by

$$\rho(\boldsymbol{a}) := \lim_{\mathrm{u}} \frac{\operatorname{rk}(a_i)}{n_i}.$$

Let further $\boldsymbol{K}$ denote the field $\prod_{\mathrm{u}} K_i$. Then $(\boldsymbol{R})_{\mathrm{u}}$ is an algebra over $\boldsymbol{K}$.

Proof. The property of being regular can be stated in first order logic, using pseudo inverses. Because matrix rings are regular rings, by Łoś' Theorem so is $\prod_{\mathrm{u}} M_{n_i}(K_i)$. Moreover the metric ultraproduct $(\boldsymbol{R})_{\mathrm{u}}$ is regular as a quotient of $\prod_{\mathrm{u}} M_{n_i}(K_i)$. Because the rings $M_{n_i}(K_i)$ are complete in the metric induced by the normalized rank, $(\boldsymbol{R})_{\mathrm{u}}$ is complete with respect to the metric induced by the rank function ρ of $(\boldsymbol{R})_{\mathrm{u}}$, by [4], Proposition 5.3. Then Theorem 4.4 ensures that $(\boldsymbol{R})_{\mathrm{u}}$ is continuous.

Let $\boldsymbol{N} := \{\boldsymbol{a} \in \boldsymbol{R} \mid \rho(\boldsymbol{a}) = 0\}$. Then $\boldsymbol{N}$ is a maximal ideal in $\boldsymbol{R}$. For, let $\boldsymbol{a} \notin \boldsymbol{N}$. Then $\rho(\boldsymbol{a}) = \delta > 0$ and hence $\frac{\operatorname{rk}(a_i)}{n_i} > \delta/2$ for u-almost all i. Thus we find $b_{i,1}, \ldots, b_{i,k}$ and $c_{i,1}, \ldots, c_{i,k}$ such that $1 = \sum_{j=1}^{k} b_{i,j} a_i c_{i,j}$ for u-almost all i, where k does only depend on δ. In the ultraproduct this means $1 = \sum_{j=1}^{k} \boldsymbol{b}_j \boldsymbol{a} \boldsymbol{c}_j$, where $\boldsymbol{b}_j$ and $\boldsymbol{c}_j$ are the elements represented by $(b_{i,j})_{i \in I}$ and $(c_{i,j})_{i \in I}$, respectively. We conclude that $\boldsymbol{N}$ is a maximal ideal and hence $(\boldsymbol{R})_{\mathrm{u}}$ is simple. Thus in particular the rank function ρ is unique by Theorem 4.5.

Since every K_i embeds into $M_{n_i}(K_i)$, $\boldsymbol{K}$ is naturally contained in $\boldsymbol{R}$. Because $\boldsymbol{K} \cap \boldsymbol{N} = \{0\}$, $\boldsymbol{K}$ embeds into $(\boldsymbol{R})_{\mathrm{u}}$. By the definition of muliplication and addition in ultraproducts, $\boldsymbol{R}$ is an algebra over $\boldsymbol{K}$, and hence so is $(\boldsymbol{R})_{\mathrm{u}}$. □

§ 2 ISOMORPHISM OF UNIVERSAL PROJECTIVELY LINEARLY SOFIC GROUPS

As is clear from the consideration of ultraproducts in Section 3, to determine whether a group has a certain approximation property is equivalent to test if this group embeds into a suitable metric ultraproduct of groups. We already hinted at the problem of finding examples of groups which cannot be approximated in matrix groups. A considerably weaker condition than embedding into a metric ultraproduct of groups is to be isomorphic to this ultraproduct. Since universal linearly sofic groups have centers isomorphic to the multiplicative group of a field, two such groups trivially cannot be isomorphic if the fields have different characteristics. For universal *projectively* linearly sofic groups this reasoning fails. The goal of this paragraph is to prove that nevertheless certain universal projectively linearly sofic groups cannot be isomorphic.

The method of proof is a modification of Ehrlich's work in [12]. Although it may seem redundant at a first glance, we give the precise formulation of the key results from [12] *and* the counterpart in our adapted setting, since we think this adds a lot to the readability of the proof.

From now on assume that R is a simple continuous regular ring. Then the center $Z(R)$ is isomorphic to a field K. Let $PR^\times$ denote the group $R^\times/Z(R^\times)$, where we reserve the symbol π for the corresponding quotient morphism and often abbreviate $\overline{g} := \pi(g)$ for $g \in R^\times$. We write elements in the lattice $L(R)$ of principal right ideals in R as $\mathfrak{a}$, $\mathfrak{b}$, etc.

The following lemma is easily verified.

Lemma 4.7 ([12], Lemma 1) *Let $2 \neq 0$ and R be a continuous regular ring. The mappings $e \mapsto 2e-1$ and $u \mapsto \frac{1}{2}(u+1)$ are mutually inverse mappings between the set of all idempotents in R and the set of all involutions in $R^\times$.*

What follows is based essentially on the above correspondence, which requires $2 \neq 0$. Consequently we will exclude the case that the field K has even characteristic from our considerations. The correspondence between idempotents and involutions can be exploited by introducing the following notions.

Let a be any element in the continuous regular ring R. Then we let

$$l^+(a) := \{x \in R \,|\, ax = x\}, \qquad l^-(a) := \{x \in R \,|\, ax = -x\}.$$

Note that $l^+(a)$ and $l^-(a)$ are by definition principal right ideals of R, and hence elements in $L(R)$. For every $\mathfrak{a} \in L(R)$ we define

$$\Delta^+(\mathfrak{a}) := \left\{ u \in R^\times \,\middle|\, u^2 = 1, l^+(u) = \mathfrak{a} \right\},$$
$$\Delta^-(\mathfrak{a}) := \left\{ u \in R^\times \,\middle|\, u^2 = 1, l^-(u) = \mathfrak{a} \right\}.$$

Sets of this form are called **Δ-sets**. So loosely speaking the further investigation is based on eigenspaces of involutions corresponding to the eigenvalues 1 and -1, and on the classification of sets of involutions with a common eigenspace. This definition yields the following:

PROPOSITION 4.8 *If $u \in R^\times$ is an involution, then*

$$l^-(u) = l^+(-u).$$

If $\mathfrak{a} \in L(R)$, then

$$\Delta^-(\mathfrak{a}) = -\Delta^+(\mathfrak{a}).$$

The next two lemmas will be useful later. Their proofs are elementary.

LEMMA 4.9 ([12], Lemma 3) *For $x \in R$ define*

$$\gamma_x : L(R) \to L(R), \qquad aR \mapsto xaR.$$

Then γ_x is an endomorphism of $L(R)$. If x is invertible, then γ_x is bijective and $\gamma_x(aR) = xax^{-1}R$.

Note that *endomorphism* in this context means *endomorphism of continuous complemented lattices*.

LEMMA 4.10 ([12], Lemma 18) *Let u be an involution and $a, b \in R$ arbitrary. Then $ua = bu$ implies $\gamma_u(l^+(a)) = l^+(b)$ and $\gamma_u(l^-(a)) = l^-(b)$.*

We will also need the following relationship between the center of a simple continuous regular ring and the center of the group of its invertible elements.

PROPOSITION 4.11 ([12], Lemma 8) *Let R be a simple continuous regular ring. Then the center of $R^\times$ equals the set of all non-zero elements in $Z(R)$.*

If in an algebra A an element $g \in A^\times$ satisfies $g^2 \in Z(A^\times)$ we call g a **projective involution** and we will use the same term for the class $\overline{g} = \pi(g) \in PA^\times$. The element g (or the corresponding class in $PA^\times$) is a projective involution of the **first kind** if g^2 is a square in $Z(A^\times)$ and of the **second kind** otherwise. Since we know that $Z(R^\times) \cong K^\times$ for the ring R, we will abuse the notation and write $g^2 = \alpha \in K^\times$ for projective involutions $g \in R^\times$.

Let $\overline{u}$ be a projective involution of the first kind in $PR^\times$. Then in the class $\overline{u}$ there are exactly two involutions. If u_0 is one of them, then the other one is $-u_0$. We define

$$L(\overline{u}) := \{l^+(u_0), l^-(u_0)\}.$$

Note that by Proposition 4.8 this definition is independent of the choice of u_0.

Let a be an element in the unital ring A. Then a is of **class** i if $(a-1)^i = 0$ and i is the smallest natural number with that property. We call $\overline{g} \in PA^\times$ of **projective class** i if it has a representative $g \in A^\times$ of class i. In a group G we write $C_G^2(M)$ for the second centralizer of the set $M \subset G$, i.e. $C_G^2(M) := C_G(C_G(M))$.

From here on we will assume that every element in the field K is a square. This has the effect that we do not have to deal with involutions of the second kind.

THEOREM 4.12 ([12], Theorem 1) *An element $s \in R^\times$ is of class 2 if and only if the following hold:*

(1) *If $t \in R^\times$, then $C_{R^\times}(t)$ properly contains $C_{R^\times}(s)$ if and only if $t \in Z(R^\times)$, i.e. $C_{R^\times}(t) = R^\times$.*

(2) *There exists an involution u such that $usu = s^{-1}$.*

(3) *There exists an element $r \in C^2_{R^\times}(u)$ such that $rsr^{-1} = s^2$.*

(4) *Except for the case when* all *elements x satisfying (1), (2) and (3) fulfill $x^3 = 1$, we have $s^3 \neq 1$.*

From this we deduce the following analogous theorem for the projective group $PR^\times$:

THEOREM 4.13 *An element $\overline{s} \in PR^\times$ is of projective class 2 if and only if the following hold:*

(1p) *If $\overline{t} \in PR^\times$, then $C_{PR^\times}(\overline{t})$ properly contains $C_{PR^\times}(\overline{s})$ if and only if $\overline{t} = 1$, i.e. $C_{PR^\times}(\overline{t}) = PR^\times$.*

(2p) *There exists a projective involution $\overline{u}$ such that $\overline{usu} = \overline{s}^{-1}$.*

(3p) *There exists an element $\overline{r} \in C^2_{PR^\times}(\overline{u})$ such that $\overline{rsr}^{-1} = \overline{s}^2$.*

(4p) *Except for the case when all elements $\overline{x}$ satisfying (1p), (2p) and (3p) fulfill $\overline{x}^3 = 1$, we have $\overline{s}^3 \neq 1$.*

Proof. Let $\overline{s}$ be any element in $PR^\times$ having properties (1p) to (4p). We aim to show that a suitable representative s of $\overline{s}$ satisfies (1) to (4) in Theorem 4.12. Let t be such that $C_{R^\times}(t)$ properly contains $C_{R^\times}(s)$, where s is *any* representative of $\overline{s}$. Then $C_{PR^\times}(\overline{t})$ properly contains $C_{PR^\times}(\overline{s})$ and hence $C_{PR^\times}(\overline{t}) = PR^\times$. Therefore $t \in Z(R^\times)$. The converse is shown similarly.

Assume now the existence of a projective involution $\overline{u}$ such that $\overline{usu} = \overline{s}^{-1}$. Then $u_0 s u_0 = \alpha^2 s^{-1}$ for appropriate α and $u_0(\alpha^{-1}s)u_0 = \alpha s^{-1}$. Since $\alpha^{-1}s$ is as good a representative for $\overline{s}$ we can without loss of generality assume $u_0 s u_0 = s^{-1}$.

Let $\overline{r}$ be such that (3p) is satisfied. Then $rsr^{-1} = \beta s^2$ for appropriate $\beta \in K^\times$ which is independent of the choice of the representative r. By left and right multiplication, respectively, we deduce $\beta = s^{-2}rsr^{-1}$ and $\beta = rsr^{-1}s^{-2}$. By conjugating the second equation with u_0,

$$\beta = r u_0 s u_0 r^{-1} u_0 s^{-2} u_0$$

follows, since in particular r commutes with u_0. Therefore an application of (2) shows

$$\beta = rs^{-1}r^{-1}s^2 = (s^{-2}rsr^{-1})^{-1} = \beta^{-1},$$

and $\beta = \pm 1$. If $\beta = -1$ holds, the equation $r(-s)r^{-1} = -rsr^{-1} = s^2 = (-s)^2$ follows. Note that still $u_0(-s)u_0 = -u_0 s u_0 = -s^{-1} = (-s)^{-1}$ and additionally (1) is true with $-s$ in place of s. Thus without loss of generality the representative s satisfies (1) to (3).

It remains to prove (4). Assume $s^3 = 1$. Then $\overline{s}^3 = 1$ and by (4p) every $\overline{x}$ with properties (1p) to (3p) is equal to 1. Let y satisfy (1) to (3). Then $\overline{y}$ satisfies (1p) to (3p) and hence $\overline{y} = 1$. This means $y^3 = \alpha \in Z(R^\times)$. By (2) there is an involution v such that $vy^3v = y^{-3}$ and hence $\alpha = \alpha^{-1}$, and $\alpha = \pm 1$ follows. If we assume $\alpha = -1$, then (3) implies the existence of t such that $ty^3t^{-1} = y^6$, leading to the contradiction $-1 = 1$. Hence $\alpha = 1$ and $y^3 = 1$. We have thus shown that s satisfies (1) to (4) and is by Theorem 4.12 of class 2.

For the converse implication let $\overline{s}$ have a representative s of class 2. Assume for (1p) that $C_{PR^\times}(\overline{t})$ properly contains $C_{PR^\times}(\overline{s})$. Since taking centralizers commutes with the quotient mapping $R^\times \to PR^\times$, $C_{R^\times}(t)$ properly contains $C_{R^\times}(s)$. By Theorem 4.12 $C_{R^\times}(t) = R^\times$ and thus $C_{PR^\times}(\overline{t}) = PR^\times$. If for the converse $C_{PR^\times}(\overline{t}) = PR^\times$ holds, we deduce by the same line of thought that $C_{PR^\times}(\overline{t})$ properly contains $C_{PR^\times}(\overline{s})$.

Conditions (2p) and (3p) follow directly from (2) and (3), respectively, when projecting $R^\times$ onto $PR^\times$.

To show (4p) assume $\overline{s}^3 = 1$. This means $s^3 = \alpha \in Z(R^\times)$ and as seen above $s^3 = 1$. Let $\overline{x}$ satisfy (1p) to (3p). Then as above x has a representative with properties (1) to (3). By property (4) $x^3 = 1$ and therefore $\overline{x}^3 = 1$. □

PROPOSITION 4.14 ([12], Proposition 6) *An element $t \in R^\times$ is of class 2 if and only if t is the product of two distinct involutions u and v such that $l^+(u) = l^+(v)$, or equivalently $l^-(u) = l^-(v)$.*

PROPOSITION 4.15 *An element $\overline{t} \in PR^\times$ is of projective class 2 if and only if $\overline{t}$ is the product of two distinct projective involutions $\overline{u}$ and $\overline{v}$ such that $L(\overline{u}) = L(\overline{v})$.*

Proof. Let $\overline{t}$ be of projective class 2 and t a representative of class 2. Then by Proposition 4.14 there are distinct involutions u and v such that $t = uv$ and $l^+(u) = l^+(v)$ as well as $l^-(u) = l^-(v)$. Note that the case $u = -v$ is clearly impossible. But then $\overline{t} = \overline{uv}$ and $L(\overline{u}) = L(\overline{v})$.

Let conversely $\overline{t}$ be the product of distinct projective involutions $\overline{u}$ and $\overline{v}$ such that $L(\overline{u}) = L(\overline{v})$. This implies $u_0v_0 = t$ for a suitable representative t, where u_0

and v_0 are distinct involutions with $l^+(u_0) = l^+(v_0)$ or $l^+(u_0) = l^-(v_0)$. Then either $u_0 v_0$ is of class 2 or $-u_0 v_0$ is of class 2. Hence $\overline{t} = \overline{\pm u_0 v_0}$ is of projective class 2. □

For a subset D of a group we write $N[D]$ for the set of all involutions in the normalizer of D.

LEMMA 4.16 ([12], Proposition 7) *If D is a Δ-set in $R^\times$, then the following properties hold:*

(1) *If $u, v, w \in D$, then $uvw = wvu \in D$.*

(2) *For every $u, v \in D$ there exists a unique $w \in R^\times$ such that $w \in D$ and $wvw = u$.*

(3) *A projective involution u is in $N[D]$ if and only if there is $v \in D$ such that $uv = vu$.*

(4) *If $t \in D^2$, then t is of projective class 1 or 2.*

The following statement is a slightly modified version of Lemma 4.16. Since in the quotient group $PR^\times$ the sets $\Delta^+(\mathfrak{a})$ and $\Delta^-(\mathfrak{a})$ are indistinguishable, we use a different formulation involving both kinds of sets at once:

LEMMA 4.17 *Let $D := \Delta^+(\mathfrak{a}) \cup \Delta^-(\mathfrak{a})$ for some $\mathfrak{a} \in L(R)$. Then D has the following properties:*

(1±) *If $u, v, w \in D$, then $uvw = wvu \in D$.*

(2±) *For every $u, v \in D$ there exist exactly two elements $w \in R^\times$ such that $w \in D$ and $wvw = \pm u$. If w is one of them the other one is $-w$.*

(3±) *A projective involution u is in $N[D]$ if and only if there is $v \in D$ such that $uv = \pm vu$.*

(4±) *If $t \in D^2$, then t or $-t$ is of class 1 or 2.*

Proof. Let $u, v, w \in D$. Then for suitable $\varepsilon_u, \varepsilon_u, \varepsilon_w \in \{\pm 1\}$ the elements $\varepsilon_u u$, $\varepsilon_v v$ and $\varepsilon_w w$ are in $\Delta^+(\mathfrak{a})$. Therefore by Lemma 4.16

$$\varepsilon_u u \cdot \varepsilon_v v \cdot \varepsilon_w w = \varepsilon_w w \cdot \varepsilon_v v \cdot \varepsilon_u u \in \varepsilon_u \varepsilon_v \varepsilon_w \Delta^+(\mathfrak{a})$$

and by cancelling the signs $uvw = wvu \in D$. Property (1±) follows.

In the same notation for $u, v \in D$ there exists a unique $w \in \Delta^+(\mathfrak{a})$ such that $\varepsilon_u u = w(\varepsilon_v v)w$, and depending on ε_u and ε_v the equation $u = \pm wvw$ follows. Moreover $-w$ is also in D and works as well. This is (2±).

The next part of the proof closely follows the proof of [12], Proposition 7. Let $u \in N[D]$ and $w \in D$. Then there exists $w' \in D$ such that $w' = uwu$ and either $l^+(w) = l^+(w')$ or $l^+(w) = l^-(w')$. We assume the latter and define

$$v := \tfrac{1}{2}(w - w') = \tfrac{1}{2}(w - uwu).$$

Then $l^+(v) = l^+(w)$ and

$$2uv = u(w - uwu) = (uwu - w)u = -2vu.$$

If $l^+(w) = l^+(w')$ defining $v := \tfrac{1}{2}(w + w')$ works.

For the converse implication let u be an involution such that $uv = \pm vu$ for some $v \in D$. We can write $v = 2f - 1$ for an appropriate idempotent f. We first assume $uv = -vu$. If $w \in D$ and $l^+(w) = l^+(v)$, then $w = 2f + 2fx(1-f) - 1$ for some element $x \in R$, by [42], Lemma 2.7. Defining $w' := -2f + 2(1-f)uxuf + 1$ yields $l^+(w') = l^-(v)$ by the same argument. Also, since $u(2f-1) = -(2f-1)u$, $uf = u - fu$ follows. We use this to derive

$$\begin{aligned} uw &= 2uf + 2ufx(1-f) - u \\ &= -2fu + 2u + (-2fu + 2u)x(1-f) - u \\ &= -2fu + u - 2fux(1-f) + 2ux(1-f). \end{aligned}$$

Then also

$$\begin{aligned}
uwu &= -2f+1-2fux(1-f)u+2ux(1-f)u \\
&= -2f+1-2fuxu+2fuxfu+2uxu-2uxfu \\
&= -2f+1-2fuxu+2fux(u-uf)+2uxu-2ux(u-uf) \\
&= -2f+1-2fuxuf+2uxuf \\
&= -2f+2(1-f)uxuf+1 = w'.
\end{aligned}$$

We proved $uw = w'u$. If $l^+(w) = l^-(v)$ the same reasoning applies after exchanging $-v$ and v. If we start from the assumption $uv = vu$ the proof is similar with $w' := 2f+2fuxu(1-f)-1$. Thus (3±) follows.

Finally, let $t \in D^2$. Then $t = uv$ and $l^+(u) = l^+(v)$ or $l^+(u) = l^-(v)$. By changing the sign if necessary t or $-t$ is of class 1 or 2. □

We will refer to a union of two corresponding Δ-sets as in the previous theorem as a $\Delta^\pm$**-set** and use the abbreviating notation

$$\Delta^\pm(\mathfrak{a}) := \Delta^+(\mathfrak{a}) \cup \Delta^-(\mathfrak{a}).$$

We call $\overline{D} \subset PR^\times$ a Λ**-set** if $\overline{D}$ is the image of a Δ-set $D \subset R^\times$ under the quotient morphism. If $D = \Delta^+(\mathfrak{a})$ or $D = \Delta^-(\mathfrak{a})$ we write

$$\Lambda(\mathfrak{a}) := \overline{\Delta^+(\mathfrak{a})} = \overline{\Delta^-(\mathfrak{a})}.$$

LEMMA 4.18 *If $\overline{D} = \Lambda(\mathfrak{a})$ is a Λ-set, then the following properties hold.*

(1p) *If $\overline{u}, \overline{v}, \overline{w} \in \overline{D}$, then $\overline{uvw} = \overline{wvu} \in \overline{D}$.*

(2p) *For every $\overline{u}, \overline{v} \in \overline{D}$ there exists a unique $\overline{w} \in PR^\times$ such that $\overline{w} \in \overline{D}$ and $\overline{wvw} = \overline{u}$.*

(3p) *A projective involution $\overline{u}$ is in $N\left[\overline{D}\right]$ if and only if there is $\overline{v} \in \overline{D}$ such that $\overline{uv} = \overline{vu}$.*

(4p) *If $\overline{t} \in \overline{D}^2$, then $\overline{t}$ is of projective class 1 or 2.*

Proof. We can without loss of generality assume that $D = \Delta^+(\mathfrak{a}) \cup \Delta^-(\mathfrak{a})$. We prove (1p). By Lemma 4.17 for $\overline{u}, \overline{v}, \overline{w} \in \overline{D}$ we have $u_0 v_0 w_0 = w_0 v_0 u_0 \in D$. Hence $\overline{uvw} = \overline{wvu} \in \overline{D}$ and the first claim follows.

By Lemma 4.17 for $\overline{u}, \overline{v} \in \overline{D}$ there exists $w \in D$, unique up to sign, such that $w v_0 w = u_0$. Therefore $\overline{wvw} = \overline{u}$, $\overline{w}$ is unique, and (2p) follows.

To prove (3p) let $\overline{u} \in N\left[\overline{D}\right]$ and $\overline{w} \in \overline{D}$. Then there is $\overline{z} \in \overline{D}$ such that $\overline{uw} = \overline{zu}$, and hence $u_0 w_0 u_0 = \alpha z_0$ for some $\alpha \in K$. By squaring this equation $\alpha = \pm 1$. Thus $u_0 \in N[D]$ and by (3±) in Lemma 4.17 there is $v \in D$ such that $u_0 v = \pm v u_0$. This implies $\overline{uv} = \overline{vu}$.

If conversely for a projective involution $\overline{u}$ there exists $\overline{v} \in \overline{D}$ such that $\overline{uv} = \overline{vu}$, then $u_0 v_0 = \alpha v_0 u_0$, where as before $\alpha = \pm 1$. Hence for $\overline{w} \in \overline{D}$ there is $z \in D$ such that $u_0 w_0 = z u_0$. This implies $\overline{uw} = \overline{zu}$.

Finally we show (4p). Let $\overline{t} \in \overline{D}^2$. Then $\overline{t} = \overline{uv}$ for $\overline{u}, \overline{v} \in \overline{D}$. Now $u_0 v_0 \in D^2$ or $-u_0 v_0 \in D^2$ is of class 1 or 2, and by (4±) in Lemma 4.17 $\overline{t}$ is of projective class 1 or 2. $\square$

LEMMA 4.19 *Let u and v be involutions in $R^\times$, $\alpha \in K$ such that $(\alpha uv - 1)^2 = 0$, i.e. αuv is of class 1 or 2. Then $\alpha = \pm 1$.*

Proof. If $uv \in K^\times$, then $\alpha uv - 1 = 0$. This implies $\alpha v = uvu$ and by squaring $\alpha^2 = 1$. Hence suppose $uv \notin K$ and $\alpha \neq \pm 1$. We evaluate $(\alpha uv - 1)^2 = 0$ to obtain

$$\alpha^2 uvuv = 2\alpha uv - 1.$$

By multiplication with v from the right $\alpha^2 uvu = 2\alpha u - v$ follows. We square this equation and use that u and v are involutions to infer

$$\alpha^4 = 4\alpha^2 - 2\alpha uv - 2\alpha vu + 1.$$

Rearranging the terms and multiplication with uv from the right yields

$$uvuv + 1 = \frac{-\alpha^4 + 4\alpha^2 + 1}{2\alpha} uv.$$

We use the relation $uvuv = 2\alpha^{-1} uv - \alpha^{-2}$ to deduce

$$\frac{-\alpha^4 + 4\alpha^2 + 1}{2\alpha} uv - 1 = 2\alpha^{-1} uv - \alpha^{-2}$$

and hence

$$vu = \frac{-\alpha^4 + 4\alpha^2 - 3}{2\alpha(1 - \alpha^{-2})}.$$

Therefore $vu \in K$ or equivalently $uv \in K$, a contradiction. Thus $\alpha = \pm 1$ as claimed. $\square$

THEOREM 4.20 *Let $\overline{D}$ be a set of projective involutions with properties (1p) to (4p) from Lemma 4.18. Then there is a set $D_\pm$ such that $\overline{D_\pm} = \overline{D}$ and $D_\pm$ satisfies properties (1±) to (4±) from Lemma 4.17*

Proof. We define $D_\pm$ to be the set of all involutions $u \in R^\times$ such that $\overline{u} \in \overline{D}$. Then for u and v in $D_\pm$ there is an involution w such that $u = \alpha w v w$, where $\alpha \in K$. By squaring this equation we obtain $\alpha^2 = 1$ and hence $\alpha = \pm 1$ follows. It is clear that $-w$ works as well as w and the pair w, $-w$ is unique. This is property (2±).

Let $t \in D_\pm^2$, i.e. $t = uv$ for involutions u and v and there is $\alpha \in K$ such that αuv is of class 1 or 2. Then $(\alpha uv - 1)^2 = 0$ and by Lemma 4.19 $\alpha = \pm 1$ follows. Therefore t is of class 1 or class 2, or $-t$ is, and we have proved (4±).

For $u, v, w \in D_\pm$ we have by Lemma 4.18, (1p) $uvw = \alpha wvu$ for some $\alpha \in K$. We rewrite this to $uv = \alpha wuvw$. Assume $(uv - 1)^2 = 0$. Then $0 = (\alpha wuvw - 1)^2$ and by multiplication with w from the left and right $(\alpha vu - 1)^2 = 0$ follows. Therefore by Lemma 4.19 $\alpha = \pm 1$. Multiplication of $(uv - 1)^2 = 0$ with vu from both sides results in the equation $(vu)^2 - 2vu + 1 = (vu - 1)^2 = 0$. The case $\alpha = -1$ gives $(vu)^2 + 2vu + 1 = 0$. These two equations together yield $4vu = 0$, which is clearly impossible. If $(-uv - 1)^2 = 0$, we replace u with $-u$ and arrive at the same conclusion. Thus $\alpha = 1$ and $uvw = wvu$. We know further that $\beta uvw \in D$ for some $\beta \in K$. As already shown $\pm(\beta uvw)w$ is of class 1 or 2 and therefore $\beta = \pm 1$. This implies $uvw \in D_\pm$ and (1±) is established.

Finally we show (3±). If $u \in N[D]$, then for $\overline{w} \in \overline{D}$ the involution w_0 is in $D_\pm$ and there exists $z \in D_\pm$ such that $uw_0 = zu$. Then $\overline{uw_0} = \overline{zu}$ and (3p) guarantees the existence of $\overline{v} \in \overline{D}$ such that $\overline{uv} = \overline{vu}$. Therefore $v_0 \in D_\pm$ and $uv_0u = \alpha v_0$ for some $\alpha \in K$. By squaring this equation $\alpha = \pm 1$ follows.

For the converse implication, let u be an involution and $v \in D$ such that $uv = \pm vu$. Then $\overline{v} \in D_\pm$ and $\overline{uv} = \overline{vu}$. Hence $\overline{u} \in N\left[\overline{D}\right]$. For $w \in D_\pm$ we have $\overline{w} \in \overline{D}$ and there is $\overline{z} \in \overline{D}$ such that $\overline{uw} = \overline{zu}$. Hence $uw = \alpha z_0 u$ for some $\alpha \in K$. Again $\alpha = \pm 1$, and since $z_0 \in D_\pm$ implies $-z_0 \in D_\pm$ as well, we conclude $u \in N[D]$. □

We need yet another result from [12]:

PROPOSITION 4.21 ([12], Proposition 3) *Let $\mathfrak{a} \in L(R)$ and $u \in R^\times$ be an involution. If $\gamma_v(\mathfrak{a}) = \mathfrak{a}$ for every involution v in the centralizer $C_{R^\times}(u)$ of u, then exactly one of $\mathfrak{a} = \{0\}$, $\mathfrak{a} = l^+(u)$, $\mathfrak{a} = l^-(u)$ or $\mathfrak{a} = R$ holds.*

Let $M \subset R$ be any subset. Then we define $l^+(M) := \bigcap_{x \in M} l^+(x)$ and $l^-(M)$ analogously. The next lemma is our counterpart to Proposition 8 in [12] and the proof is almost the same.

LEMMA 4.22 *Let D be a set of involutions with properties (1±) to (4±) and $D^{(2)}$ the set of all elements of class 1 or 2 in D^2. Then $D^{(2)}$ is an abelian subgroup of $R^\times$. If D contains at least three elements, then*

$$D \subset \Delta^\pm(l^+(D^{(2)})).$$

Proof. Let s and t be in $D^{(2)}$, where $s = xy$, $t = uv$ for $x, y, u, v \in D$. Then by (1±)

$$st = xyuv = uyxv = uvxy = ts$$

follows.

Now assume that D contains at least three involutions. Let s, t be in $D^{(2)}$ and write $n := s - 1$, $m := t - 1$. We then have $n^2 = m^2 = 0$ and since $D^{(2)}$ is abelian, $(1+n)(1+m) = (1+m)(1+n)$, which implies $nm = mn$. Because $D^{(2)}$ is a group, st is of class 2, which means $((1+n)(1+m)-1)^2 = 0$. We deduce

$$\begin{aligned} 0 &= (nm + n + m)^2 \\ &= nmnm + nm(n+m) + (n+m)nm + (n+m)^2 \\ &= n^2m^2 + 2n^2m + 2nm^2 + n^2 + m^2 + 2nm = 2nm. \end{aligned}$$

This implies $sm = (1+n)m = m$ and $m \in l^+(s)$. Since s and t were chosen arbitrarily, $t-1 \in l^+(D^{(2)})$ for all $t \in D^{(2)}$. We assumed the existence of at least three distinct involutions in D, u, v and w say. If $uv = 1$, then $u = v$ and otherwise exactly one of uv and $-uv$ is of class 2. Therefore one element t in $D^{(2)}$ is not 1, and consequently $l^+(D^{(2)}) \neq \{0\}$. Since there are more than two elements in D, by (4±) 1 is not the only element in $D^{(2)}$ and therefore $l^+(D^{(2)}) \neq R$.

Now consider an involution $u \in N[D]$. Then trivially $u \in N\left[D^2\right]$. If $t \in D^{(2)}$, there exists $t_u \in D^2$ such that $ut = t_u u$. We compute $0 = (t-1)^2 = (ut_u u - 1)^2$, whence $0 = (t_u - 1)^2$, $t_u \in D^{(2)}$ and even $u \in N\left[D^{(2)}\right]$. By Lemma 4.10 $\gamma_u(l^+(t)) = l^+(t_u)$ for every $u \in N[D]$ and $t \in D^{(2)}$, where $ut = t_u u$. Now

$$\begin{aligned}\gamma_u(l^+(D^{(2)})) &= \gamma_u\left(\bigcap_{t\in D^{(2)}} l^+(t)\right) = \bigcap_{t\in D^{(2)}} \gamma_u(l^+(t)) \\ &= \bigcap_{t\in D^{(2)}} l^+(t_u) \subset l^+(D^{(2)}).\end{aligned}$$

Because γ_u is a lattice automorphism by Lemma 4.9, in fact $\gamma_u(l^+(D^{(2)})) = l^+(D^{(2)})$ holds, for any $u \in N[D]$.

If $v \in D$ and w is an involution in the centralizer $C_{R^\times}(v)$, then by property (3±) w is in $N[D]$. Hence $l^+(D^{(2)})$ is invariant under γ_w for every such $w \in C_{R^\times}(v)$. By Proposition 4.21 $l^+(D^{(2)})$ equals either $\{0\}$, $l^+(v)$, $l^-(v)$ or R. Since we have shown that $\{0\}$ and R are no possible candidates, $l^+(D^{(2)}) = l^+(v)$ or $l^+(D^{(2)}) = l^-(v)$. We assume the former and now let $u \in D$ be arbitrary. Then there is $w \in D$ such that $v = \pm wuw$. By Lemma 4.10 one of

$$l^+(u) = \gamma_w(l^+(v)) = \gamma_w(l^+(D^{(2)})), \qquad l^+(-u) = \gamma_w(l^+(D^{(2)}))$$

holds. Because $w \in D$ implies $w \in N[D]$, we also have $\gamma_w(l^+(D^{(2)})) = l^+(D^{(2)})$, and $l^+(u) = l^+(D^{(2)})$ or $l^+(-u) = l^+(D^{(2)})$ follows. Hence

$$u \in \Delta^+(l^+(D^{(2)})) \cup \Delta^-(l^+(D^{(2)})).$$

The case $l^+(D^{(2)}) = l^-(v)$ is similar and $D \subset \Delta^\pm(l^+(D^{(2)})$ follows as claimed. □

The proof of the next lemma is clear.

LEMMA 4.23 *The mapping* $\Delta^{\pm}(\mathfrak{a}) \mapsto \Lambda(\mathfrak{a})$ *is injective.*

The following theorem including proof is modeled after [12], Theorem 2.

THEOREM 4.24 *The set* $D \subset G$ *is a* $\Delta^{\pm}$*-set if and only if it is a maximal set with properties* (1±) *to* (4±) *from Lemma* 4.17.

The set $\overline{D} \subset PR^{\times}$ *is a* Λ*-set if and only if it is a maximal set with properties* (1p) *to* (4p) *from Lemma* 4.18.

Proof. Let $D := \Delta^{\pm}(\mathfrak{a})$ for some $\mathfrak{a} \in L(R)$. By Lemma 4.17 D has properties (1±) to (4±) defined therein. Assume there is a strictly larger set E with these properties, and define $\mathfrak{b} := l^{+}(E^{(2)})$. Then

$$l^{+}(E^{(2)}) = \mathfrak{b} \subset \mathfrak{a} = l^{+}(D^{(2)}).$$

Since E is larger than D and D contains at least two elements, there are at least three elements in E. By Lemma 4.22 $E \subset \Delta^{\pm}(\mathfrak{b})$. For $u \in E \setminus D$ we know $l^{+}(u) = \mathfrak{b}$ or $l^{-}(u) = \mathfrak{b}$. In the first case $l^{+}(u) = \mathfrak{b} \subset \mathfrak{a} = l^{+}(D^{(2)})$ and hence the contradiction $u \in D$ follows. Otherwise we replace u with $-u$ for the same contradiction. Hence follows maximality of D.

Let conversely D be a maximal set of involutions satisfying (1±) to (4±). It is clear that $u \in D$ implies $-u \in D$ and thus D cannot consist of only one element. We show that $D = \{u, -u\}$ implies $u = \pm 1$. If we assume the converse, then $l^{+}(u) \neq \{0\}$ and $l^{+}(u) \neq R$. In an irreducible continuous geometry the only elements with unique complement are 0 and 1, i.e. $\{0\}$ and R in our setting. (Confer [32], Chapter I, § 3.) Therefore there exists $\mathfrak{b}$ such that $\mathfrak{b} \cap l^{+}(u) = \{0\}$ and $\mathfrak{b} \vee l^{+}(u) = R$. By Lemma 2 in [12] there is an involution $v \neq u$ such that $l^{+}(v) = l^{+}(u)$. This means $v \in \Delta^{+}(l^{+}(u))$ and D is properly contained in $\Delta^{+}(l^{+}(u))$, contradicting its maximality. Therefore $u = \pm 1$ and D consists of all involutions in R. If there are more than two elements in D, then by Lemma 4.22 $D = \Delta^{\pm}(\mathfrak{a})$ for $\mathfrak{a} = l^{+}(D^{(2)})$.

By Lemma 4.23 the claim concerning Λ-sets follows. □

We have now reached the important point where Λ-sets are completely characterized within the group $PR^{\times}$. The next step will exploit the close relation between Λ-sets and elements in $L(R)$.

LEMMA 4.25 *If $D = \Delta^{\pm}(\mathfrak{a})$, then $D^{(2)} = \Delta^{+}(\mathfrak{a})^2$.*

P r o o f. The inclusion from right to left is clear. Let $t \in D^{(2)}$. Then $t = uv$, where u and v are in D. If u and v are from the same Δ-set, then t is of class 1 or 2 by Proposition 4.14. Otherwise $-t$ is of class 1 or 2. As seen in the proof of Theorem 4.20 t and $-t$ cannot be both of class 1 or 2. Hence $t \in D^{(2)}$ implies $t \in \Delta^{+}(\mathfrak{a})^2$. □

THEOREM 4.26 *Let $\mathfrak{a}, \mathfrak{b}, \mathfrak{c} \in L(R)$. Then $\mathfrak{a}$ and $\mathfrak{b}$ are both different from $\{0\}$ and R and $\mathfrak{c}$ is between $\mathfrak{a}$ and $\mathfrak{b}$ in the ordering of $L(R)$, if and only if*

(1) $\Lambda(\mathfrak{a}) \subset N[\Lambda(\mathfrak{b})]$ *or* $\Lambda(\mathfrak{b}) \subset N[\Lambda(\mathfrak{a})]$, *and*

(2) $\{1\} \subset \Lambda(\mathfrak{a})^2 \cap \Lambda(\mathfrak{b})^2 \subset \Lambda(\mathfrak{c})^2$.

P r o o f. Theorem 3 in [12] states that $\mathfrak{a}$ and $\mathfrak{b}$ do both not equal $\{0\}$ or R and $\mathfrak{c}$ is between $\mathfrak{a}$ and $\mathfrak{b}$, if and only if

(1) $\Delta^{+}(\mathfrak{a}) \subset N[\Delta^{+}(\mathfrak{b})]$ or $\Delta^{+}(\mathfrak{b}) \subset N[\Delta^{+}(\mathfrak{a})]$, and

(2) $\{1\} \subset \Delta^{+}(\mathfrak{a})^2 \cap \Delta^{+}(\mathfrak{b})^2 \subset \Delta^{+}(\mathfrak{c})^2$.

Now we use Lemma 4.23, Lemma 4.25 and the fact that $\overline{u} \in N[\Lambda(\mathfrak{a})]$ if and only if $\pm u_0 \in N[\Delta^{+}(\mathfrak{a})]$. □

THEOREM 4.27 *The mappings*

$$\mathfrak{a} \mapsto \Lambda(\mathfrak{a}), \qquad \Lambda(\mathfrak{a}) \mapsto l^{+}(\Delta^{\pm}(\mathfrak{a})^{(2)})$$

are mutually inverse and define a one-to-one correspondence between the set of all non-zero elements in $L(R)$ and the set of all $\Delta^{\pm}$-sets in $R^{\times}$.

P r o o f. This follows using Theorem 5 in [12], which states that

$$\mathfrak{a} \mapsto \{\Delta^{+}(\mathfrak{a}), \Delta^{-}(\mathfrak{a})\}, \qquad \{D, -D\} \mapsto l^{+}(D^2)$$

are inverse mappings defining a one-to-one correspondence between $L(R) \setminus \{\{0\}\}$ and the set of all pairs of Δ-sets to a common $\mathfrak{a}$ in $R^{\times}$. By Lemma 4.23 and Lemma 4.25 the claim follows. □

A **dual isomorphism** φ of two lattices L_1 and L_2 is a bijective mapping $L_1 \to L_2$ such that $\varphi(a \wedge b) = \varphi(a) \vee \varphi(b)$ and $\varphi(a \vee b) = \varphi(a) \wedge \varphi(b)$ for all $a, b \in L_1$. An **anti-isomorphism** $\psi\colon R_1 \to R_2$ of two rings R_1 and R_2 is a bijective mapping preserving the additive structure and satisfying $\psi(ab) = \psi(b)\psi(a)$ for all $a, b \in R_1$.

THEOREM 4.28 *Let $2 \neq 0$ and R_1, R_2 be two simple continuous regular rings, every element in the centers of which is a square. Then R_1 and R_2 are isomorphic or anti-isomorphic if and only if the projective groups $PR_1^\times$ and $PR_2^\times$ are isomorphic.*

P r o o f. If R_1 and R_2 are isomorphic or anti-isomorphic clearly the groups $PR_1^\times$ and $PR_2^\times$ are isomorphic.

Now assume that $\varphi\colon PR_1^\times \to PR_2^\times$ is an isomorphism. By Theorem 4.24 $\varphi(\Lambda(\mathfrak{a}))$ is a Λ-set in $PR_2^\times$ for all $\mathfrak{a} \in L(R_1)$. We write $\mathfrak{a}_\varphi$ for the unique element in $L(R_2)$ such that

$$\varphi(\Lambda(\mathfrak{a})) = \Lambda(\mathfrak{a}_\varphi).$$

By Theorem 4.27, and since by [12], Theorem 4 $\Lambda(\mathfrak{a}) = \{1\}$ if and only if $\mathfrak{a} = \{0\}$ or $\mathfrak{a} = R$, φ induces a bijection

$$\psi\colon L(R_1) \setminus \{\{0\}, R_1\} \to L(R_2) \setminus \{\{0\}, R_2\}, \qquad \mathfrak{a} \mapsto l^+(\Delta^\pm(\mathfrak{a}_\varphi)^{(2)}).$$

By Theorem 4.26 the mapping ψ is order preserving or order reversing. Now Lemma 25 in [12] ensures that ψ is induced by an isomorphism or dual isomorphism of $L(R_1)$ and $L(R_2)$. Then by [42], Part II, Theorem 4.3 R_1 and R_2 are isomorphic or anti-isomorphic. □

THEOREM 4.29 *Let $\Gamma_1 := \left(\prod \mathrm{PGL}_{n_i}(K_i)\right)_\mathfrak{u}$ and $\Gamma_2 := \left(\prod \mathrm{PGL}_{m_i}(F_i)\right)_\mathfrak{u}$ be universal projectively linearly sofic groups, i.e. the ultraproducts in question are defined with respect to the Jordan length. Let $\boldsymbol{K} := \prod_\mathfrak{u} K_i$ and $\boldsymbol{F} := \prod_\mathfrak{u} F_i$. If $\boldsymbol{K}$ and $\boldsymbol{F}$ have different characteristics, both not equal to 2, and every element in both fields is a square, then Γ_1 and Γ_2 are not isomorphic.*

P r o o f. Theorem 4.6 shows that $\boldsymbol{R}_1 := \left(M_{n_i}(K_i)\right)_\mathfrak{u}$ and $\boldsymbol{R}_2 := \left(M_{m_i}(F_i)\right)_\mathfrak{u}$ are simple continuous regular rings. By Theorem 2.22 the groups $\boldsymbol{R}_1^\times$ and $\boldsymbol{R}_2^\times$ are isomorphic to the metric ultraproducts of the groups $\mathrm{GL}_{n_i}(K_i)$ and $\mathrm{GL}_{m_i}(F_i)$, respectively, defined using the rank length. Then $P\boldsymbol{R}_j^\times \cong \Gamma_j$ for $j = 1, 2$. By Proposition 4.2 the centers of $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ are fields and by Theorem 4.6 they contain $\boldsymbol{K}$ and $\boldsymbol{F}$, respectively. Thus the respective centers share the characteristic of $\boldsymbol{K}$ and $\boldsymbol{F}$, and all elements in the centers are squares. By Theorem 4.28 isomorphism of Γ_1 and Γ_2 would imply that $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$ are isomorphic or anti-isomorphic, in which case their centers would be isomorphic. Since the centers were then isomorphic to fields of different characteristics, this is impossible. $\square$

Part II

Normal subgroups in ultraproducts of compact simple groups

5 ULTRAPRODUCTS OF FINITE SIMPLE GROUPS

§ 1 FINITE SIMPLE GROUPS OF LIE TYPE

The Classification of Finite Simple Groups states that all finite simple groups appear in the exhaustive list of

(1) abelian finite simple groups (i.e. cyclic groups of prime order),

(2) alternating groups A_n,

(3) groups of Lie type,

(4) the 26 sporadic groups.

There are different ways to introduce groups of Lie type. The approach justifying the name and deriving their classification is by simple Lie groups and can be found in [10]: The classification of complex simple Lie algebras states that there are four main families A_n, B_n, C_n and D_n of complex simple Lie algebras indexed by $n \in \mathbb{N}$, and five exceptional ones, G_2, F_4, E_6, E_7 and E_8. Every finite simple group of Lie type then derives from one of these Lie algebras. This method is called the *Chevalley Construction*. The groups obtained from Lie algebras of type A_n, B_n, C_n and D_n are called **classical groups**. The others are the **exceptional groups**. The subscript in the identifiers of simple Lie algebras is the rank of the corresponding Lie algebra. We will associate the same number, the **rank**, with every simple group of Lie type. This implies in particular that classical groups can have arbitrarily large rank, while the rank of the exceptional groups is bounded. Note that this introduction is rather incomplete and serves mainly to distinguish classical and exceptional groups and to give a motivation for the notion of rank. For reasons which will become clear in § 4 the classical groups need the most attention and we neglect a more detailed description of the exceptional groups.

Although we will also deal with simple Lie groups later, we will stick to the geometric approach to handle the classical groups of Lie type, which is best suited for our purposes and has the advantage of requiring the least theoretical buildup. In the following we therefore describe the classical finite simple groups of Lie type as quotients of isometry groups of finite-dimensional vector spaces over finite fields with bilinear or Hermitian forms. All details can be found in [24]. Let in the following V denote a finite-dimensional vector space. If V is equipped with a bilinear or Hermitian form $(\ ,\ )$ we write $\mathrm{GI}(V)$ for the group of all linear isometries $V \to V$, i.e. the group of all linear mappings $\varphi\colon V \to V$ satisfying

$$(\varphi(v), \varphi(w)) = (v, w)$$

for all $v, w \in V$. The subgroup of all isometries of determinant 1 is denoted by $\mathrm{SI}(V)$.

The first class of groups to consider are the **projective special linear groups** $\mathrm{PSL}(V) := \mathrm{SL}(V)/Z(\mathrm{SL}(V))$. We can interprete $\mathrm{SL}(V)$ to be the group of isometries preserving the trivial, i.e. identically zero bilinear form on V.

If V is endowed with a symplectic bilinear form, the group of isometries is the **symplectic group**, denoted by $\mathrm{Sp}(V) := SI(V) = GI(V)$. The **projective symplectic group** is $\mathrm{PSp}(V)$, the quotient of $\mathrm{Sp}(V)$ by its center.

The **unitary group** $\mathrm{U}(V)$ is the group of isometries with respect to a Hermitian form on V. We also have the **special unitary group** $\mathrm{SU}(V) := SI(V)$ in this context and the **projective special unitary group** $\mathrm{PSU}(V)$, the quotient of $\mathrm{SU}(V)$ by its center. Note that for unitary groups the vector space V is always defined over a field $\mathbb{F}_{q'}$, where $q' = q^2$ is a square.

The last groups to consider are a bit more complicated. Assume first that the field of definition of V has odd characteristic. Then $\mathrm{GO}(V) := \mathrm{SI}(V)$ and $\mathrm{SO}(V) := \mathrm{SI}(V)$ for a symmetric bilinear form on V, and these are called the **orthogonal group** or **special orthogonal group**, respectively. In characteristic 2 defining $\mathrm{SO}(V)$ as the subgroup of elements of determinant 1 does not work. Instead $\mathrm{SO}(V)$ is defined as the kernel of the Dickson invariant. (Confer [24], pp.129-131.) Let $\Omega(V) := \mathrm{GO}(V)' = \mathrm{SO}(V)'$, the commutator subgroup of $\mathrm{SO}(V)$. Then the **projective orthogonal group** is $\mathrm{PSO}(V) := \Omega(V)/Z(\Omega(V))$.

Now the groups $\mathrm{PSL}(V)$, $\mathrm{PSp}(V)$, $\mathrm{PSU}(V)$ and $\mathrm{PSO}(V)$, where V is a vector space over a finite field, are classical groups of Lie type. They are simple with a few exceptions occuring when $\dim(V) \leq 5$. A perfect group which is simple modulo its center is called **quasisimple**. Hence the groups $\mathrm{SL}(V)$, $\mathrm{Sp}(V)$, $\mathrm{SU}(V)$ and $\Omega(V)$ are the quasisimple counterparts to the simple groups we are interested in.

If V has dimension m and is defined over $\mathbb{F}_q$ we will write $\mathrm{Sp}_m(q)$, $\mathrm{SU}_m(q)$ and so on, as done before in the case of general and special linear groups. To link the geometric approach to the Lie theoretic approach we note that $\mathrm{PSL}(V)$ and $\mathrm{PSU}(V)$ derive from Lie algebras of type A_n, $\mathrm{PSp}(V)$ from a Lie algebra of type C_n, and $\mathrm{PSO}(V)$ from a Lie algebra of type B_n if $\dim(V)$ is odd, and of type D_n if $\dim(V)$ is even.

From now on we will always mean finite simple *non-abelian* groups whenever we simply write finite simple groups.

§ 2 METRIC ULTRAPRODUCTS OF FINITE SIMPLE GROUPS

As already seen in Theorem 2.26, the length functions ℓ_c and ℓ_J are asymptotically equivalent in the classes of general and special linear groups. We shall extend this result to symplectic, orthogonal and unitary groups.

THEOREM 5.1 *Let $\mathcal{G}$ be the class of classical finite simple groups of Lie type. Then the pseudo length functions ℓ_c and ℓ_J are asymptotically equivalent in $\mathcal{G}$.*

Proof. The proof is almost the same as the proof of Theorem 2.26 and 2.28, except for the references from [31]. If $G \in \mathcal{G}$ and $g \in G$ we write m for the maximal number of Jordan blocks of size 1 to a common eigenvalue in the Jordan decomposition of g. It is sufficient to work in quasisimple groups, because the center of a quasisimple group defined over the field $\mathbb{F}_q$ has at most q elements, which is small enough to be disregarded in the estimates below. Hence we can assume that G is a symplectic, orthogonal or special linear group, and we write $n = \dim(V)$, where V is the vector space G is naturally acting on. As in the proof of Theorem 2.28 we have $m_g - m \leq \frac{1}{2}(n-m)$. We use Lemma 6.3 in [31] if G is a symplectic or orthogonal group and ibid., Section 7 for unitary groups to see that there is a constant c such that

$$\frac{\log|C(g)|}{\log|G|} \leq c\frac{n^2 - m^2}{n^2}.$$

This implies

$$\ell_{\mathrm{c}}(g) \leq c_0 \frac{n - m_g}{n} = c_0 \ell_{\mathrm{J}}(g),$$

for another constant c_0 depending only on c. Therefore ℓ_{c} is asymptotically bounded by ℓ_{J}.

Now by Lemma 6.4 in [31] for symplectic or orthogonal G and ibid., Section 7 for unitary G, there is a constant c_1 such that for $k \geq \frac{c_1 n}{n-m}$ and $g \in G \setminus \{1\}$ we have $C(g)^k = G$. The proof now continues exactly as the proof of Theorem 2.26 to show that ℓ_{J} is asymptotically bounded by ℓ_{c}. □

PROPOSITION 5.2 *Let $(G)_{\mathrm{u}}$ be a metric ultraproduct of classical finite simple groups G_i of Lie type, defined over the field $\mathbb{F}_q$. Then $(G)_{\mathrm{u}}$ is q-sofic.*

P r o o f. We can assume without loss of generality that every G_i is quasisimple. Then G_i is a subgroup of $\mathrm{GL}_{n_i}(q)$ for appropriate n_i and the rank length on G_i is the same as on $\mathrm{GL}_{n_i}(q)$. By Theorem 3.7 $(G)_{\mathrm{u}}$ has the $\mathcal{GL}(q)$-approximation property with respect to the rank length. This implies the claim. □

The converse is also true in the following sense.

THEOREM 5.3 *Let G_i be finite quasisimple groups of Lie type of rank r_i over the field $\mathbb{F}_q$, such that $\lim_{\mathrm{u}} r_i = \infty$. Then the universal q-sofic group $\left(\prod \mathrm{GL}_{r_i}(q)\right)_{\mathrm{u}}$ embeds into the metric ultraproduct $\left(\prod G_i\right)_{\mathrm{u}}$.*

P r o o f. Let G be a symplectic, orthogonal or unitary group $\mathrm{SI}(V)'$ of rank r, where $\dim(V) = n$. Then V has a totally isotropic subspace U of dimension r, i.e. $(v, w) = 0$ for all $v, w \in U$. Moreover there is a complementary subspace U' of the same dimension. Now $\mathrm{GL}(U)$ embeds into G by simultaneous action of $g \in \mathrm{GL}(U)$ on U and on U', subject to the condition that the action is an isometry. The action on the complement of $U \oplus U'$ is trivial. If $g \in \mathrm{GL}(U)$ we write g' for the element in G resulting from this embedding.

Now if $g \in \mathrm{GL}(U) \setminus \{1\}$ satisfies $m = \dim\ker(1-g)$, then

$$\dim\ker(1-g') \in \{2m, 2m+1, 2m+2\}$$

depending on the geometry of V, where $\dim(V) \in \{2r, 2r+1, 2r+2\}$, respectively. Hence

$$\begin{aligned}\ell_{\mathrm{r}}(g') &= 1 - \frac{\dim\ker(1-g')}{n} \leq 1 - \frac{2m}{2r+2} \\ &\leq 2\left(1 - \frac{m}{r}\right) = 2\ell_{\mathrm{r}}(g),\end{aligned}$$

for r large enough. Conversely also

$$\ell_{\mathrm{r}}(g) = 1 - \frac{m}{r} \leq 1 - \frac{2m+2}{2r} \leq 2\ell_{\mathrm{r}}(g)$$

holds. □

§ 3 SUBGROUPS OF FINITE SIMPLE GROUPS

We need some basic geometric lemmas to prepare what follows. We use the symbol $\ominus$ to denote the orthogonal direct sum. The dual space of a vector space V is denoted by V^*. If V is equipped with a bilinear or Hermitian form $(\ ,\ )$, then the orthogonal complement $W^\perp$ of a subspace W is the set $\{v \in V \mid \forall w \in W : (v,w) = 0\}$. The **radical** $\mathrm{rad}(W)$ of W is the subspace $W \cap W^\perp$, and W is called **degenerate** if $\mathrm{rad}(W) \neq \{0\}$.

The next lemma is essentially Corollary 2.3 in [24].

LEMMA 5.4 *Let V be a finite-dimensional vector space with non-degenerate bilinear or Hermitian form $(\ ,\ )$ and W a subspace. If $\varphi \in W^*$, then there is $v \in V$ such that for all $w \in W$ the equation $(w,v) = \varphi(w)$ holds.*

Proof. Because V is non-degenerate, $\{0\} = \mathrm{rad}(V) = \ker(v \mapsto (w \mapsto (w,v)))$. Hence for all $\varphi \in V^*$ there is $v \in V$ such that $\varphi(w) = (w,v)$ for all $w \in V$. Let now W be a subspace with basis $w_1, \dots, w_k$, which extends to a basis $w_1, \dots, w_n$ of V. Extend $\varphi \in W^*$ to a linear form $\tilde{\varphi}$ on V by $\tilde{\varphi}(w_i) = 0$ if $i > k$ and find v such that $\varphi(w) = (w, v)$ for all w in W. □

LEMMA 5.5 *Let V be a finite-dimensional vector space with non-degenerate bilinear or Hermitian form $(\ ,\ )$ and W a subspace. Let R be the radical of W and W' a complement of R in W. Then there is a subspace W'' of V which satisfies $\dim(W'') = \dim(R)$ and $\left(W'' \oplus W^\perp\right) \ominus W' = V$. In particular W' and $U := W'' \oplus W^\perp$ are non-degenerate.*

Proof. We use Lemma 5.4. Let $w_1, \dots, w_r$ be a basis of R and $w_1, \dots, w_k$ an extension to a basis of $W + W^\perp$. For $r = 0$ there is nothing to show since $W \oplus W^\perp = V$. Now assume $r \geq 1$ and define $\varphi_1 \in \left(W + W^\perp\right)^*$ by $\varphi_1(w_1) = 1$ and $\varphi_1(w_i) = 0$ otherwise. Then there is $v_1 \in V$ such that $(w_1, v_1) = 1$ and $v \perp \langle w_2, \dots, w_k \rangle$. Now $\dim\left(\operatorname{rad}\left(W + W^\perp + \langle v_1 \rangle\right)\right) = r - 1$ and we can proceed inductively defining $\varphi_l \in \left(W + W^\perp + \langle v_1, \dots, v_{l-1} \rangle\right)^*$ by $\varphi_l(w_l) = 1$ and $\varphi_l(w_i) = 0$, $\varphi_l(v_i) = 0$ for the remaining basis vectors. In the end this gives us $v_1, \dots, v_r$ such that $W'' := \langle v_1, \dots, v_r \rangle$ has the claimed properties. □

LEMMA 5.6 *Let V be a finite-dimensional vector space over a field K with bilinear or Hermitian form $B = (\ ,\ )$. We exclude the case that $\operatorname{char}(K) = 2$ and B symmetric. Let U be a non-degenerate subspace of V. Then the subgroup*

$$H := \left\{ g \in \mathrm{SI}(V) \,\middle|\, g|U^\perp = \mathrm{id}_{U^\perp} \right\}$$

of $\mathrm{SI}(V)$ is isomorphic to $\mathrm{SI}(U)$.

Proof. Let g be in $\mathrm{SI}(U)$. We define $\varphi(g) := g \oplus \mathrm{id}_{U^\perp}$. Obviously $\varphi(g) \in \mathrm{GL}(V)$ and φ is a homomorphism. If v, w are in V and decompose as $v = v_U + v_{U^\perp}$, $w = w_U + w_{U^\perp}$ with respect to the direct sum $V = U \oplus U^\perp$, then

$$\begin{aligned}(\varphi(g)(v), \varphi(g)(w)) &= (g(v_U), g(w_U)) + (v_{U^\perp}, w_{U^\perp}) \\ &= (v_U, w_U) + (v_{U^\perp}, w_{U^\perp}) = (v, w),\end{aligned}$$

whence $\varphi(g) \in \mathrm{GI}(V)$. If $\det(g) = 1$, obviously also $\det(\varphi(g)) = 1$ and $\varphi(g) \in \mathrm{SI}(V)$. Because elements in H stabilize $U^\perp$, $g|U$ is guaranteed to stabilize U. Now the inverse mapping of φ is easily verified to be $g \mapsto g|U$. □

Let (,) be a symmetric bilinear form on a vector space V over a field K. Then the associated **quadratic form** Q is defined by

$$Q(v) := (v, v).$$

Let g be in $\mathrm{GO}(V)$ and assume $\mathrm{char}(K) \neq 2$. Then g can be written in different ways as a product of reflections, $g = s_{u_1} \cdot \ldots \cdot s_{u_j}$ say. We write $K^{\times 2}$ for the set of all squares in $K^\times$. Then

$$\vartheta\colon \mathrm{GO}(V) \to K^\times / K^{\times 2}, \qquad v \mapsto Q(s_{u_1}) \cdot \ldots \cdot Q(s_{u_j}) K^{\times 2}$$

is a well defined map by [24], Chapter 9, pp. 75, 76. It is called the **spinor norm**. If K has characteristic 2, then the spinor norm is defined similarly with orthogonal transvections instead of reflections. (Confer [24], p. 137.)

LEMMA 5.7 *Let V be a finite-dimensional vector space over a field K of odd characteristic with non-degenerate symmetric bilinear form (,). Let U be a non-degenerate subspace. Then the subgroup*

$$H := \left\{ g \in \Omega(V) \,\middle|\, g|U^\perp = \mathrm{id}_{U^\perp} \right\}$$

of $\Omega(V)$ is isomorphic to $\Omega(U)$.

Proof. Let $g \in H$. Because $g|U \in \mathrm{GO}(U)$, it can can be written as a product of reflections $s_{u_1}, \ldots, s_{u_k}$ in $\mathrm{GO}(U)$, where the reflection is along the hyperplane $\langle u_i \rangle^\perp$ (and u_i non-degenerate). In particular $u_i \in U$ for all i. Each reflection s_{u_i} is given explicitly by the expression

$$s_{u_i}(w) = w - Q(u_i)^{-1}(w, u_i)u_i,$$

where Q is the associated quadratic form. From orthogonality we deduce that $s_{u_i}^V := s_{u_i} \oplus \mathrm{id}_{U^\perp}$ is a reflection in $\mathrm{GO}(V)$.

By Lemma 5.6 we know that $g|U \in \mathrm{SO}(V)$. By [24], Theorem 9.7

$$\Omega(V) = \mathrm{SO}(V) \cap \ker(\vartheta),$$

where ϑ is the spinor norm. We see

$$\begin{aligned}1 = \vartheta(g) &= \vartheta(s_{u_1}^V \dots s_{u_k}^V) \\ &= Q(u_1) \cdot \ldots \cdot Q(u_k) K^{\times 2} \\ &= \vartheta(s_{u_1} \dots s_{u_k}) = \vartheta(g|U)\end{aligned}$$

and conclude $g|U \in \Omega(U)$. The claim follows. □

LEMMA 5.8 *Let V be a vector space of dimension n over a perfect (finite) field K of characteristic 2. Let Q be a regular, i.e.* $\dim(\mathrm{rad}(V)) \in \{0, 1\}$, *quadratic form on V and $B = (,)$ the associated bilinear form. Let U be a subspace of V such that $Q|U$ is regular and*

$$H := \left\{ g \in \Omega(V) \,\middle|\, g|U^\perp = \mathrm{id}_{U^\perp} \right\}.$$

Then H is isomorphic to $\Omega(U)$.

Proof. We note that H is isomorphic to a subgroup of $\mathrm{GO}(U)$ by $g \mapsto g|U$. We divide the proof according to the dimension of the radical of V and assume first that V is defective, i.e. $\mathrm{rad}(V) \neq \{0\}$. Then by Theorem 14.2 in [24] $\mathrm{GO}(V)$ is isomorphic to $\mathrm{Sp}(V_1)$ for a complement V_1 of $\mathrm{rad}(V)$ and the action of $\mathrm{GO}(V)$ on $\mathrm{rad}(V)$ is trivial. We see that whether U is defective or not, the proof of Lemma 5.6 applies.

Now assume that V is non-defective. By [24], Proposition 14.23

$$\Omega(V) = \mathrm{SO}(V) \cap \ker(\vartheta),$$

where ϑ is the spinor norm. Note that $\mathrm{SO}(V)$ is the kernel of the Dickson invariant $\delta : \mathrm{GO}(V) \to \mathbb{F}_2$, or equivalently the subgroup of all products of an even number of orthogonal transvections. Now if $g \in H$, then $g|U \in \mathrm{GO}(U)$ and hence is a product of transvections $t_{u_1}, \dots, t_{u_k}$. (Confer [24], Theorem 14.16.) Each orthogonal transvection is described explicitly by

$$t_{u_i}(w) = w + Q(u_i)^{-1}(w, u_i) u_i.$$

We implicitly used that none of the vectors u_i is singular, i.e. $Q(u_i) \neq 0$ for all i. By extending $g|U$ to the whole of V as in Lemma 5.7, k is necessarily even. The proof now continues as in Lemma 5.7. □

§ 4 THE ORDERING OF THE SET OF NORMAL SUBGROUPS IN ULTRAPRODUCTS OF FINITE SIMPLE GROUPS

The investigation of normal subgroups of ultraproducts of finite simple groups starts with the following theorem.

THEOREM 5.9 ([18], Theorem 1.1) *Let $\mathfrak{u}$ be a non-principal ultrafilter in the natural numbers. Then the set of normal subgroups of the ultraproduct $\prod_{\mathfrak{u}} A_n$ of alternating groups is linearly ordered.*

Another formulation of this statement can be found in [1] as Theorem 3.

We will generalize Theorem 5.9 to ultraproducts of arbitrary finite simple non-abelian groups. Therefore we recall the discussion at the end of Section 3, § 1: Given an ultraproduct G of arbitrary finite simple non-abelian groups, the chosen ultrafilter "decides" whether G is isomorphic to an ultraproduct of e.g. groups of bounded or unbounded rank, permutation groups or groups of Lie type, or in the case of groups of Lie type of large rank which type of classical group they belong to. As a consequence there will be in particular no further treatment of sporadic groups, since an ultraproduct of simple groups of bounded finite cardinality is again finite and simple.

Note also that we could replace the groups A_n in Theorem 5.9 by groups A_{n_i}, where i is from an arbitrary index set. Then either $n_i \leq k$ for $\mathfrak{u}$-almost all i and a constant k, and the ultraproduct itself is isomorphic to one A_{n_j}, where $n_j \leq k$. Or $\lim_{\mathfrak{u}} n_i = \infty$ and in this case Theorem 5.9 can be proved exactly as in [18].

The next theorem, the main result in [31], will be used several times and we include the statement here for the sake of convenience.

THEOREM 5.10 ([31], Theorem 1.1) *There is a universal constant c such that whenever G is a finite non-abelian simple group and $1 \neq g \in G$, then*

$$m \geq c \frac{\log|G|}{\log|C(g)|}$$

implies $C(g)^m = G$.

We obtain the following proposition for groups of Lie type almost instantly.

PROPOSITION 5.11 *Let $\mathcal{G} = \{G_i \mid i \in I\}$ be a family of finite non-abelian simple groups. Then the metric ultraproduct group $\left(\prod_{i\in I} G_i\right)_{\mathfrak{u}}$ is simple.*

P r o o f. We show that if $\ell_c(\boldsymbol{g}) = \varepsilon > 0$ for $\boldsymbol{g} \in \boldsymbol{G}$, then already $N(\boldsymbol{g}) = \boldsymbol{G}$. By Theorem 5.10 and our assumption $\frac{\log|G_{\omega(i)}|}{\log|C(g_i)|} \leq K$ $[\mathfrak{u}]$. Hence for $m \geq cK$, $C(g_i)^m = G_{\omega(i)}$ $[\mathfrak{u}]$ or equivalently $C(\boldsymbol{g})^m = \boldsymbol{G}$. We conclude that the set of all infinitesimal elements in $\boldsymbol{G}$ is a maximal normal subgroup and thus $\boldsymbol{G}$ divided by this subgroup is simple. □

In fact the converse is also true. If a quotient of a direct product of finite simple non-abelian groups is simple, then it is a quotient as in the preceding theorem for some choice of ultrafilter. This was proved in [33], Proposition 3.

THEOREM 5.12 *Let G_i be finite simple groups of Lie type for all $i \in I$, and $\mathfrak{u}$ a non-principal ultrafilter in I. If $\boldsymbol{G} = \prod_{\mathfrak{u}} G_i$ and the rank of the groups G_i is bounded, then $\boldsymbol{G}$ is simple.*

P r o o f. Suppose that the rank of the groups in question is bounded by N. Let $1 \neq g \in G_i$. Using the constant c of Theorem 5.10 we see that for

$$m \geq \frac{c \log|G_i|}{\log|C(g)|}$$

already $C(g)^m = G_i$. When G_i is a group over the field $\mathbb{F}_q$, its order is at most $q^{c'N^2}$ for a constant c'. On the other hand a non-trivial conjugacy class in G_i has at least q elements. Hence it suffices for m to be larger than $c'N^2$ to ensure $C(g)^m = G_i$ for any i.

If we choose $1 \neq \boldsymbol{g} \in \boldsymbol{G}$ arbitrarily, then $C(g_i)^m = G_i$ for $\mathfrak{u}$-almost all i. Hence $C(\boldsymbol{g})^m = \boldsymbol{G}$ and consequently $N(\boldsymbol{g}) = \boldsymbol{G}$. Therefore $\boldsymbol{G}$ contains no proper normal subgroups, whence it is simple. □

With the preceding theorem the treatment of exceptional groups of Lie type is complete, because they are of bounded rank.

We take Theorems 5.9 and 5.12 as a motivation to prove the following more general Theorem 5.13. In the proof we follow a similar route as the authors of [18] in the proof of their Theorem 1.1.

THEOREM 5.13 *Let G_i be finite simple groups of Lie type for all $i \in I$. If $\mathbf{G} := \prod_{\mathfrak{u}} G_i$, then the set $\mathfrak{N}$ of normal subgroups of $\mathbf{G}$ is linearly ordered.*

In view of Theorem 5.12 we only need to take care of classical groups of unbounded rank.

First consider the general situation that we are given an ultraproduct $\mathbf{G} = \prod_{\mathfrak{u}} G_i$ of arbitrary groups G_i with length function ℓ_i, such that $\operatorname{diam}(G_i)$ is bounded. We define an ordering of the non-trivial elements of $\mathbf{G}$ by $\mathbf{g} \preceq \mathbf{h}$ if

$$\lim_{\mathfrak{u}} \frac{\ell_i(g_i)}{\ell_i(h_i)} < \infty.$$

LEMMA 5.14 *Let $\mathbf{g}$ and $\mathbf{h}$ be non-identity elements of the ultraproduct $\mathbf{G}$ of groups G_i. Then $\mathbf{g} \in N(\mathbf{h})$ implies $\mathbf{g} \preceq \mathbf{h}$.*

P r o o f. If $\mathbf{g} \in N(\mathbf{h})$ there is some integer k such that $\mathbf{g}$ is a product of k conjugates of $\mathbf{h}^{\pm 1}$. Therefore g_i is a product of k conjugates of $h_i^{\pm 1}$ for $\mathfrak{u}$-almost all i. By the properties of invariant length functions

$$\ell_i(g_i) \leq k\ell_i(h_i^{\pm 1}) = k\ell_i(h_i) \; [\mathfrak{u}].$$

Hence

$$\lim_{\mathfrak{u}} \frac{\ell_i(g_i)}{\ell_i(h_i)} \leq k$$

follows and we are done. □

The plan is to show that for finite simple groups of Lie type and the Jordan length the converse of the previous lemma is true. The following statement is a summary of results from [31].

LEMMA 5.15 *Let $G = \mathrm{SI}(V)$ be a quasisimple group of Lie type, where V has dimension n and $g \in G \setminus Z(G)$. There is a constant c, independent of G and g, such that $C(g)^k = G$ for all $k \geq \frac{cn}{n-m_g}$.*

Proof. Let m denote the maximal number of Jordan blocks of size 1 in the Jordan decomposition of $g \in G$. Then it is clear that $m \leq m_g$, and $\frac{cn}{n-m_g} \geq \frac{cn}{n-m}$ follows. The conclusion can be derived using m instead of m_g by [31]: For special linear groups use Lemma 5.4, for symplectic and orthogonal groups Lemma 6.4 and for unitary groups Section 7 ibid. $\square$

LEMMA 5.16 *Let $\boldsymbol{G} = \prod_{\mathfrak{u}} G_i$ be an ultraproduct of finite simple groups of Lie type equipped with the Jordan length. Then $\boldsymbol{g} \preceq \boldsymbol{h}$ implies $\boldsymbol{g} \in N(\boldsymbol{h})$ for all non-trivial elements $\boldsymbol{g}, \boldsymbol{h} \in \boldsymbol{G}$.*

Proof. Note that we can safely neglect exceptional groups, since these are of bounded rank and hence dealt with in Theorem 5.12. More generally we assume that $G_i = \mathrm{SI}(V_i)$, where V is a vector space of dimension n_i and $\lim_{\mathfrak{u}} n_i = \infty$. These groups are only quasisimple, but working with the Jordan length will produce the same result in the ultraproduct. By the hypothesis there is a natural number k such that $\frac{n_i - m_{g_i}}{n_i - m_{h_i}} \leq k$ for $\mathfrak{u}$-almost all i.

Let $G = \mathrm{SI}(V)$, where V has dimension n. We can exclude the case when the characteristic of the field of definition is 2 and V is a defective quadratic space from the following considerations, since under that assumptions $G = \mathrm{GO}(V)$ is isomorphic to a symplectic group. Assume that $\frac{n-m_g}{n-m_h} \leq k$ for some elements $g, h \in G \setminus Z(G)$ such that $n - m_g = \mathrm{rk}(1-g)$ and $n - m_h = \mathrm{rk}(1-h)$, that is their rank length and Jordan length are the same. We define $W := \ker(1-g) \cap \ker(1-h)$. If W' is a complement of $\mathrm{rad}(W)$ in W, following Lemma 5.5 there is subspace W'' such that $U := W'' \oplus W^{\perp}$ is non-degenerate and $W' = U^{\perp}$. Obviously g and h act as the identity on $U^{\perp}$. Then $g|U$ and $h|U$ are in $H := \mathrm{SI}(U)$. We compute

$$\dim(W^{\perp}) = n - \dim(W) \leq n - (m_g + m_h - n) = (n - m_g) + (n - m_h)$$

and

$$\dim(\mathrm{rad}(W)) \leq n - \dim(W) \leq (n - m_g) + (n - m_h).$$

This together with the introductory remarks implies

$$\begin{aligned}\dim(U) &= \dim(W^{\perp}) + \dim(\mathrm{rad}(W))\\ &\leq 2(n-m_g) + 2(n-m_h)\\ &\leq (2k+2)(n-m_h).\end{aligned}$$

Therefore the Jordan length of $h|U$ estimates as

$$\ell_{\mathrm{J}}(h|U) = \frac{\dim(U) - (\dim(U) - (n-m_h))}{\dim(U)} = \frac{n-m_h}{\dim(U)} \geq \frac{1}{2k+2}.$$

By Lemma 5.15 there is a constant c, independent of the hypotheses, such that $((h|U)^H)^m = H$ for $m \geq c(2k+2)$ and consequently $g|U$ is a product of m conjugates of $h|U$ inside H. As in Lemma 5.6 we extend the elements occuring in this product to elements in G, thereby extending $g|U$ to g and $h|U$ to h. Thus the conclusion remains true in G and also when returning attention to the finite simple group $G/Z(G)$.

Because the prototype $G/Z(G)$ was independent of $i \in I$ and the hypotheses did hold for almost all i, g_i is a product of $m \geq c(2k+2)$ conjugates of h_i in G_i for almost all i. Hence

$$\boldsymbol{g} \in C(\boldsymbol{h})^m \subset N(\boldsymbol{h}),$$

which we had to prove. □

COROLLARY 5.17 *If $\boldsymbol{g}$ and $\boldsymbol{h}$ are non-identity elements in $\boldsymbol{G}$, the statements $\boldsymbol{g} \in N(\boldsymbol{h})$ and $\boldsymbol{g} \preceq \boldsymbol{h}$ are equivalent.*

The last preparation we need is Lemma 2.2 in [18], which for the sake of completeness we cite with proof:

LEMMA 5.18 *Let G be any group. Then the set of normal subgroups of G is linearly ordered by inclusion if and only if the set of normal closures of non-identity elements in G is.*

Proof. The first implication is trivial. For the converse assume that N and M are normal subgroups of G such that $N \not\subset M$. Let $g \in N \setminus M$ and observe that necessarily $N(g) \not\subset N(h)$ for all $h \in M$. Thus $N(h) \subset N(g)$ for all $h \in M$, and $M \subset N$ follows. □

P r o o f of Theorem 5.13. We define a quasiorder on the set $L := \prod_{\mathfrak{u}}[n]$ by $a \preceq b$ if $a_n \leq b_n$ for $\mathfrak{u}$-almost all n. We let furthermore $a \equiv b$, whenever

$$0 < \lim_{\mathfrak{u}} \frac{a_n}{b_n} < \infty.$$

Therefore $\equiv$ is a convex equivalence relation, i.e. $a \preceq b \preceq c$ and $a \equiv c$ imply $a \equiv b \equiv c$. The quotient space $L/\equiv$ is totally ordered, because the real numbers are.

By the foregoing considerations, culminating in Corollary 5.17, the set of normal closures of elements in G is order isomorphic to $L/\equiv$. Lemma 5.18 shows that the set of normal subgroups of G is linearly ordered by inclusion if and only if the set of normal closures of elements of G is. Now Theorem 5.13 follows. □

6 Ultraproducts of compact connected simple Lie groups

§ 1 Compact simple Lie groups

The explanation of Lie group terms we give here is anything but complete or precise. Nevertheless everything we need to know is standard Lie theory as can be found in most books on this topic. We will use the references [8], [39] and [28], where the first is good for a comprehensive introduction, the second concentrates on compact Lie groups and the last one is recommended due to its very detailed treatment.

The classification of compact connected simple Lie groups implies that these are

(1) special unitary groups $\mathrm{PSU}(n+1) \coloneqq \mathrm{PSU}_{n+1}(\mathbb{C})$,

(2) orthogonal groups $\mathrm{PSO}(2n+1) \coloneqq \mathrm{PSO}_{2n+1}(\mathbb{C})$,

(3) symplectic groups $\mathrm{PSp}(2n) \coloneqq \mathrm{PSp}_{2n}(\mathbb{C})$,

(4) orthogonal groups $\mathrm{PSO}(2n) \coloneqq \mathrm{PSO}_{2n}(\mathbb{C})$,

(5) the **exceptional** groups G_2, F_4, E_6, E_7 and E_8,

and the list is exhaustive. The first four items on the list comprise the so called **classical** groups.

If G is a compact connected Lie group, then we call G **quasisimple** if it is perfect and simple modulo its center. In contrast to certain customs in Lie theory we will use the term *simple* only for groups which are actually simple as abstract groups. The classical simple groups as introduced above arise as central quotients of $\mathrm{SU}(n+1) \coloneqq \mathrm{SU}_n(\mathbb{C})$, $\mathrm{SO}(2n+1) \coloneqq \mathrm{SO}_{2n+1}(\mathbb{C})$, $\mathrm{Sp}(2n) \coloneqq \mathrm{Sp}_{2n}(\mathbb{C})$ and $\mathrm{SO}(2n) \coloneqq \mathrm{SO}_{2n}(\mathbb{C})$, respectively. Note that not all of these groups are simply connected. However to each of them corresponds a simply connected group with the same central quotient and the same Lie algebra, obtained as the universal covering group.

The compact connected quasisimple Lie groups are classified by their Lie algebra, which is of type A_n for $\mathrm{SU}(n+1)$ and $\mathrm{PSU}(n+1)$, of type B_n for $\mathrm{SO}(2n+1)$ and $\mathrm{PSO}(2n+1)$, of type C_n for $\mathrm{Sp}(2n)$ and $\mathrm{PSp}(2n)$, and of type D_n for $\mathrm{SO}(2n)$ and $\mathrm{PSO}(2n)$. This classification is also the reason for dividing the orthogonal groups into two classes. The Lie algebras of the exceptional groups are of type G_2, F_4, E_6, E_7 and E_8, where it should be clear which Lie algebra belongs to which group.

As a decisive means of distinction between the exceptional groups and the classical groups we need the notion of rank. The Maximal Torus Theorem ([8], Theorem 16.5) states that every compact connected Lie group G has a maximal closed connected abelian subgroup T, called a **maximal torus**. Moreover every element in G is conjugate to an element in T. The dimension of T as a manifold is called the **rank** of G. If T' is another maximal torus, then T and T' are conjugate, and hence the rank is well defined. The classical groups in the above list have rank n. The rank of the exceptional groups is the number in the index. If N denotes the normalizer of T, then N/T is called the **Weyl group** of G. It can be realized as a subgroup of G.

Corresponding to every maximal torus T there is a subset Φ of the character group of T, consisting of the **roots** of G with respect to T. Since the character group of T is isomorphic to $\mathbb{Z}^r$ if G has rank r, one can think of roots as elements of $\mathbb{R}^r$. The set of roots seen as vectors in euclidean space forms a so called **root system**, and the lattice (where the term lattice here means a discrete subgroup) spanned by the roots is called the **root lattice**. Every root system contains a set of **fundamental roots** $\{\beta_1, \dots, \beta_r\}$. Note that the number of fundamental roots is always the same as the rank of G and the corresponding root vectors are linearly independent. When we consider the β_i as homomorphisms $T \to S^1$, we know that

$$\bigcap_{i=1}^{r} \ker \beta_i = Z(G).$$

Chapters 19 and 20 in [8] present an introduction to root systems.

We shall make use of the fundamental roots in explicit calculations and therefore take a closer look at fundamental roots of classical Lie groups. In § 2 and § 5 we are going to use fixed unitary representations of different types of Lie groups which we will refer to as standard representations. (Confer [8], Chapter 20 for these representations.) Since a maximal torus of $\mathrm{U}(n) := \mathrm{U}_n(\mathbb{C})$ is the subgroup T of diagonal matrices, if we embed G in $\mathrm{U}(n)$ we can find a maximal torus of G inside T.

We use the obvious embedding of $\mathrm{SU}(n+1)$ in $\mathrm{U}(n+1)$. Then we obtain the maximal torus of diagonal matrices $t = \mathrm{diag}(t_1, \ldots, t_{n+1})$ of determinant 1 as the basis of our considerations. Corresponding to this torus the fundamental roots of $\mathrm{SU}(n+1)$ are given by

$$\beta_i(t) = t_i t_{i+1}^{-1}$$

for $i = 1 \ldots n$.

We have $\mathrm{Sp}(2n)$ realized inside $\mathrm{U}(2n)$ as matrices of the form

$$\begin{pmatrix} a & -\overline{b} \\ b & \overline{a} \end{pmatrix}$$

with complex $n \times n$-matrices a, b. Then the maximal torus of our choice consists of elements $t = \mathrm{diag}(t_1, \ldots, t_n, t_1^{-1}, \ldots, t_n^{-1})$. The fundamental roots are

$$\beta_i(t) = t_i t_{i+1}^{-1}, \qquad \beta_n(t) = t_n^2,$$

where i is between 1 and $n-1$. The first $n-1$ fundamental roots are called **short roots** and β_n is a **long root**.

The orthogonal matrices $\mathrm{SO}(2n+1)$ embed into $\mathrm{U}(2n+1)$ as the subgroup of all $g \in \mathrm{U}(2n+1)$ such that $gJg^{\mathrm{T}} = J$, where g^{T} is the transpose of g and

$$J := \begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix}.$$

Then a maximal torus consists of elements $t = \mathrm{diag}(t_1, \ldots, t_n, 1, t_n^{-1}, \ldots, t_1^{-1})$ and the fundamental roots are

$$\beta_i(t) = t_i t_{i+1}^{-1}, \qquad \beta_n(t) = t_n,$$

for $i = 1 \dots n-1$. Here the long roots are $\beta_1, \dots, \beta_{n-1}$ and β_n is a short root.

The situation is similar for SO($2n$), except that torus elements are of the form $t = \operatorname{diag}(t_1, \dots, t_n, t_n^{-1}, \dots, t_1^{-1})$ (where the 1 in the middle is missing) and the fundamental roots are

$$\beta_i(t) = t_i t_{i+1}^{-1}, \qquad \beta_n(t) = t_{n-1} t_n.$$

Note that for SU($n+1$) and SO($2n$) we will not need the distinction between long and short roots and hence omit it. (In fact all fundamental roots have the same length for these groups.)

For every character α there is a cocharacter $\eta_\alpha : S^1 \to T$ such that $\alpha(\eta_\alpha(z)) = z^2$ for all $z \in S^1$. For every pair of roots α, $-\alpha$ in Φ there is a homomorphism $\varphi_\alpha : \mathrm{SU}(2) \to G$ such that η_α is the restriction of φ_α to the subgroup of diagonal matrices in SU(2). (Confer [8], Chapter 20.) For every root α we define associated subgroups. We write $T_\alpha \coloneqq \{g \in T \mid \alpha(g) = 1\} \subset T$, i.e. T_α equals the kernel of α. We let $S_\alpha = S_{-\alpha}$ denote the image of $\mathrm{SU}(2) \subset G$ under φ_α. Then S_α commutes elementwise with T_α and T is contained in the central product $S_\alpha T_\alpha$. At last we define the one parameter torus H_α as the image of the cocharacter η_α, thus $H_\alpha \subset S_\alpha$. For fundamental roots β_i we use the self-explanatory shorthand notation T_i, S_i and H_i. Then T equals the direct product $H_1 H_2 \dots H_r$.

We will need a last technical preparation. If Φ is a root system corresponding to a group G of rank r and we interprete the roots as vectors in $\mathbb{R}^r$, the set of all ω such that

$$2\frac{(\omega, \alpha)}{(\alpha, \alpha)} \in \mathbb{Z}$$

for all $\alpha \in \Phi$, where $(\ ,\ )$ is the euclidean scalar product, is a lattice in $\mathbb{R}^r$. It is called the **weight lattice** and its elements **weights**. The root lattice is contained in the weight lattice. The weights $\omega_1, \dots, \omega_r$ satisfying

$$2\frac{(\omega_i, \beta_j)}{(\beta_j, \beta_j)} = \delta_{ij}$$

are called **fundamental dominant weights**. To each fundamental dominant weight corresponds an irreducible representation of G, the **fundamental representation**. In the context of the construction of this representation the used weight is called a **highest weight**. Chapters 20, 24 and 25 in [8] present all the facts about weights we use.

For use in § 2 only, we choose standard representations for the exceptional groups. For E_7, E_8, F_4 and G_2 let δ be the smallest fundamental representation and for E_6 the second smallest fundamental representation, of maximal dimension 351 in the case of E_6. As the standard representation we use

$$\delta' \colon g \mapsto \delta(g) \oplus \overline{\delta}(g),$$

where $\overline{\delta}$ results from the composition of δ and complex conjugation of matrix entries.

In the following G will always denote a quasisimple compact connected Lie group and T a maximal torus in G.

§ 2 LENGTH IN COMPACT CONNECTED QUASISIMPLE LIE GROUPS

In this paragraph we shall examine two different ways of measuring distances in certain compact connected quasisimple Lie groups.

PROPOSITION 6.1 *Let* $||\cdot||$ *be a matrix norm which evaluates to* 1 *on the group* $\mathrm{U}(n)$*. Then*

$$\ell(g) := \frac{1}{2n}||1-g||$$

defines an invariant length function on $\mathrm{U}(n)$*.*

P r o o f. This is a corollary of Proposition 2.20. □

We will need the length function ℓ_1, obtained by the preceding proposition when using the norm $||A|| := \sum_{i=1}^{n} |a_{ij}|$, where $A = (a_{ij})_{i,j=1\ldots n}$ is a complex matrix.

A complex number z in S^1 can be written uniquely as $e^{i\vartheta}$, where $\vartheta \in]-\pi,\pi]$. We call $l(z) := |\vartheta|$ the **angle** of z. Now we define

$$\lambda(g) := \frac{1}{\pi r}\sum_{i=1}^{r} l(\beta_i(g))$$

for all $g \in T$. Proposition 5.11 in [34], the proof of which is spread over Subsection 5.5 ibid., includes the following result:

PROPOSITION 6.2 *The function $\lambda\colon T \to \mathbb{R}$ is an invariant pseudo length function on T, and $\lambda(g) = 0$ if and only if $g \in Z(G)$.*

Unfortunately λ is only defined on the torus T of a Lie group G, but not on G itself. Since every element of G is conjugate to one in T, a natural idea would be to extend λ to G by defining $\lambda(g) := \lambda(t)$, where $g \in G$ is conjugate to $t \in T$. The main problem is that λ is not invariant under conjugation and hence the extension to G would not be well defined. Consider for example SU(3) and the torus element

$$t = \operatorname{diag}(t_1, t_2, t_3) = \operatorname{diag}(e^{\frac{1}{3}\pi \mathrm{i}}, e^{\frac{4}{3}\pi \mathrm{i}}, e^{\frac{1}{3}\pi \mathrm{i}}).$$

Then

$$\lambda(t) = \frac{1}{2\pi}(l(\beta_1(t)) + l(\beta_2(t))) = \frac{1}{2\pi}(l(t_1 t_2^{-1}) + l(t_2 t_3^{-1})) = \frac{\pi + \pi}{2\pi}$$

holds. But t is conjugate to $t' = \operatorname{diag}(e^{\frac{1}{3}\pi \mathrm{i}}, e^{\frac{1}{3}\pi \mathrm{i}}, e^{\frac{4}{3}\pi \mathrm{i}})$ in SU(3) and we compute $\lambda(t') = \frac{1}{2} \neq 1 = \lambda(t)$.

To achieve invariance under conjugation we replace λ with the following function. We define

$$\tilde{\lambda}(g) := \sup_{t \in C(g) \cap T} \lambda(t).$$

Since every element in G is conjugate to one in T, this is a reasonable definition. This new function has the advantage that, in contrast to λ, it is invariant under conjugation. Moreover, as can be expected it sometimes takes considerably larger values than λ, a fact we shall exploit later.

We are now in control of the functions $\tilde{\lambda}$ and ℓ_1, the former of which is well adapted to the Lie group structure while the latter is a length function and moreover has features which will become apparent in § 5. We will proceed by showing that these functions are related closely enough to work with both and exploit the respective advantages.

LEMMA 6.3 *Let ϑ be a real number in the interval $[-\pi,\pi]$. Then*

$$\tfrac{1}{4}\sqrt{2}\sqrt{1-\cos\vartheta} \le \frac{|\vartheta|}{\pi} \le \tfrac{1}{2}\sqrt{2}\sqrt{1-\cos\vartheta}.$$

P r o o f. First assume $\vartheta \in [0,\pi]$. Let $f(\vartheta) := 8\frac{\vartheta^2}{\pi^2} - 1 + \cos\vartheta$. Then $f(0) = 0$ and $f'(\vartheta) = \frac{16}{\pi^2}\vartheta - \sin\vartheta \ge \vartheta - \sin\vartheta \ge 0$. This implies $f(\vartheta) \ge 0$ and consequently $\frac{1}{8}(1-\cos\vartheta) \le \frac{\vartheta^2}{\pi^2}$. By taking the square root the first estimate follows. If $g(\vartheta) := 1 - \cos\vartheta - 2\frac{\vartheta^2}{\pi^2}$, we have $g'(\vartheta) = \sin\vartheta - \frac{4}{\pi^2}\vartheta$. Moreover $g'(0) = 0$ and $g''(\vartheta) = \cos\vartheta - \frac{4}{\pi^2}$. Hence the monotone decreasing function g'' has a unique zero in $[0,\pi]$, starting from $g''(0) > 0$. Therefore g' has a unique maximum in $[0,\pi]$. This and $g(0) = 0$ imply that g is positive near 0 and it can have at most two zeros in $[0,\pi]$. Since these zeros are clearly 0 and π, g is non-negative on $[0,\pi]$, which proves the second estimate. By symmetry of the cosine the claim holds for negative ϑ. □

LEMMA 6.4 *There are $L, L' > 0$ such that the following holds: Let G be a classical compact connected quasisimple Lie group of rank r contained in a unitary group $\mathrm{U}(n)$ by the respective standard embedding. We write*

$$\ell_1'(g) := \inf_{z \in Z(U(n))} \frac{n}{r}\ell_1(zg)$$

for elements $g \in G$. Then for any g in G

$$L^{-1}\ell_1'(g) \le \tilde{\lambda}(g) \le L'\ell_1'(g).$$

P r o o f. In $\mathrm{U}(n)$ we can write

$$\ell_1(g) = \frac{1}{2n}\sum_{i=1}^{n}|1-\mu_i|,$$

where μ_i are the eigenvalues of g. For real ϑ we have

$$|1-e^{i\vartheta}| = \sqrt{2}\sqrt{1-\cos\vartheta}.$$

Then by Lemma 6.3

$$\frac{l(e^{i\vartheta})}{\pi} \le \tfrac{1}{2}|1-e^{i\vartheta}| \le 2\frac{l(e^{i\vartheta})}{\pi}.$$

Let first G be equal to $\mathrm{SU}(n+1)$. For diagonal elements $t = \mathrm{diag}(t_1, \dots, t_{n+1})$ in the torus T of diagonal matrices of determinant 1 we have $\beta_i(t) = t_i t_{i+1}^{-1}$. Therefore (abusing notation to apply β_i to elements not in $\mathrm{SU}(n+1)$) we have $\lambda(t) = \lambda(zt)$ for any central element $z = \mathrm{diag}(z, \dots, z) \in Z(U(n+1))$. Because ℓ_1 is a length function on S^1, $|1 - xy^{-1}| \leq |1-x| + |1-y|$ for all $x, y \in S^1$, and hence

$$|1 - zt_i(zt_{i+1})^{-1}| \leq |1 - zt_i| + |1 - zt_{i+1}|.$$

Therefore the estimate

$$\begin{aligned}\lambda(t) &= \inf_{z \in Z(\mathrm{U}(n+1))} \lambda(zt) \\ &= \inf_{z \in Z(\mathrm{U}(n+1))} \frac{1}{\pi n} \sum_{i=1}^{n} l(zt_i(zt_{i+1})^{-1})) \\ &\leq \inf_{z \in Z(\mathrm{U}(n+1))} \frac{1}{n} \sum_{i=1}^{n+1} |1 - zt_i| = 2\ell_1'(t)\end{aligned}$$

follows. Then, taking the supremum over all $t \in C(g) \cap T$, also $\tilde{\lambda}(g) \leq 2\ell_1'(g)$ holds for any $g \in \mathrm{SU}(n+1)$. Given t we can reorder its diagonal entries by conjugation with a generalized permutation matrix (i.e. a permutation matrix with entries in ± 1 and determinant 1), such that without loss of generality $l(t_1 t_2^{-1})$ is maximal among all possible values $l(t_i t_j^{-1})$. Proceeding from this point we can achieve inductively that $l(t_i t_{i+1}^{-1}) \geq l(t_i t_j^{-1})$ for all $j > i$. This yields $l(t_i t_{n+1}^{-1}) \leq l(t_i t_{i+1}^{-1})$ for all $i = 1 \dots n$. Now

$$\begin{aligned}\ell_1'(t) = \inf_{z \in Z(\mathrm{U}(n+1))} \frac{1}{2n} \sum_{i=1}^{n+1} |1 - t_i z| &\leq \inf_{z \in Z(\mathrm{U}(n+1))} \frac{2}{\pi n} \sum_{i=1}^{n+1} l(t_i z) \\ \leq \frac{2}{\pi n} \sum_{i=1}^{n+1} l(t_i t_{n+1}^{-1}) &\leq \frac{2}{\pi n} \sum_{i=1}^{n} l(t_i t_{i+1}^{-1}) = 2\lambda(t)\end{aligned}$$

follows. Therefore $\ell_1'(g) \leq 2\tilde{\lambda}(g)$ holds for any $g \in \mathrm{SU}(n+1)$.

Now consider $\mathrm{SO}(2n+1) \subset \mathrm{U}(2n+1)$. An element in the maximal torus of $\mathrm{SO}(2n+1)$ then has the form $t = \mathrm{diag}(t_1, \dots, t_n, 1, t_n^{-1}, \dots, t_1^{-1})$. The characters corresponding to fundamental roots are given by $\beta_i(t) = t_i t_{i+1}^{-1}$ for $i = 1 \dots n-1$ and $\beta_n(t) = t_n$. We want to proceed as in the case of $\mathrm{SU}(n+1)$ but have to take care of the different fundamental root. Every root has to be estimated twice, since ℓ_1 counts both $|1-t_i|$ and $|1-t_i^{-1}|$, and so we shall use

$$\begin{aligned}2|1-zt_i(zt_{i+1})^{-1}| &= |1-zt_i(zt_{i+1})^{-1}|+|1-zt_i^{-1}z^{-1}t_{i+1}|\\ &\le |1-zt_i|+|1-zt_{i+1}|+|1-zt_i^{-1}|+|1-zt_{i+1}^{-1}|.\end{aligned}$$

Noting that $t_{n+1}=1$ and $t_i^{-1}=t_{2n+2-i}$, we see

$$\begin{aligned}\lambda(t) &= \frac{1}{\pi n}\sum_{i=1}^{n} l(t_i t_{i+1}^{-1}) \le \frac{1}{2n}\sum_{i=1}^{n}|1-zt_i(zt_{i+1})^{-1}|\\ &\le \frac{1}{2n}\sum_{i=1}^{n}|1-zt_i|+|1-zt_i^{-1}|+\frac{1}{4n}(|1-z|+|1-z^{-1}|)\\ &= \frac{1}{2n}\sum_{i=1}^{2n+1}|1-zt_i| = \frac{2n+1}{n}\ell_1(zt).\end{aligned}$$

By taking the infimum over all $z\in Z(U(2n+1))$, $\lambda(t)\le \ell_1'(t)$ follows, and because this is independent of the ordering of the t_i, also $\tilde{\lambda}(g)\le \ell_1'(g)$ for arbitrary $g\in SO(2n+1)$. To reorder the entries of $\operatorname{diag}(t_1,\dots,t_n,1,t_n^{-1},\dots,t_1^{-1})$ by conjugation with a permutation matrix there are the possibilities to permute t_i and t_i^{-1}, and to permute the first n entries, which entails corresponding permutation of the last n. Hence without loss of generality we can assume $l(t_1t_2^{-1})$ maximal among all $l(t_it_j^{\pm 1})$ and $l(t_it_{i+1}^{-1})\ge (t_it_j^{\pm 1})$ for $j>i$. Then

$$\begin{aligned}\ell_1'(t) &= \inf_{z\in Z(U(2n+1))}\frac{1}{2n}\sum_{i=1}^{n}|1-t_iz|+|1-t_i^{-1}z|+\frac{1}{2n}|1-z|\\ &\le \inf_{z\in Z(U(2n+1))}\frac{2}{\pi n}\sum_{i=1}^{n} l(t_iz)+l(t_i^{-1}z)+\frac{2}{\pi n}l(z)\\ &\le \frac{2}{\pi n}\sum_{i=1}^{n-1} l(t_it_n)+l(t_i^{-1}t_n)+\frac{6}{\pi n}l(t_n)\\ &\le \frac{4}{\pi n}\sum_{i=1}^{n-1} l(t_it_{i+1}^{-1})+\frac{6}{\pi n}l(t_n)\le 6\lambda(t)\end{aligned}$$

implies $\ell_1'(g)\le 6\tilde{\lambda}(g)$ for any $g\in SO(2n+1)$.

We continue right away with $SO(2n)$, where the characters evaluate as $\beta_i(t)=t_it_{i+1}^{-1}$ if $i<n$ and $\beta_n(t)=t_{n-1}t_n$. The computations to obtain $\tilde{\lambda}(g)\le 2\ell_1'(g)$

are similar to the case of SO$(2n+1)$, in fact

$$\begin{aligned}
\lambda(t) &= \frac{1}{\pi n}\sum_{i=1}^{n-1} l(t_i t_{i+1}^{-1}) + \frac{1}{\pi n} l(t_{n-1}t_n) \\
&\leq \frac{1}{2n}\sum_{i=1}^{n-1} |1 - z t_i (z t_{i+1})^{-1}| + \frac{1}{2n}|1 - z t_{n-1} z^{-1} t_n| \\
&\leq \frac{1}{2n}\sum_{i=1}^{n-1} |1 - z t_i| + |1 - z t_i^{-1}| + \frac{1}{4n}(|1 - z t_n| + |1 - z t_n^{-1}|) \\
&\quad + \frac{1}{2n}(|1 - z t_{n-1}| + |1 - z t_n^{-1}|) \\
&\leq \frac{1}{n}\sum_{i=1}^{2n} |1 - z t_i| = \frac{2\cdot 2n}{n} \ell_1(zt).
\end{aligned}$$

We can permute entries of diagonal elements as in the case of SO$(2n+1)$, with the restriction of performing only an even number of exchanges $t_i \mapsto t_i^{-1}$. But regardless of whether we have $t_n^{\pm 1}$ in the right place, the estimate

$$\begin{aligned}
\ell_1'(t) &= \inf_{z\in \mathrm{Z}(\mathrm{U}(2n))} \frac{1}{2n}\sum_{i=1}^{n} |1 - t_i z| + |1 - t_i^{-1} z| \\
&\leq \inf_{z\in \mathrm{Z}(\mathrm{U}(2n))} \frac{2}{\pi n}\sum_{i=1}^{n} l(t_i z) + l(t_i^{-1} z) \\
&\leq \frac{2}{\pi n}\sum_{i=1}^{n-1} l(t_i t_n) + l(t_i^{-1} t_n) + \frac{2}{\pi n} l(t_n^2) \\
&\leq \frac{4}{\pi n}\sum_{i=1}^{n-1} l(t_i t_{i+1}^{-1}) + \frac{2}{\pi n}(l(t_n t_{n-1}^{-1}) + l(t_n t_{n-1})) \\
&\leq \frac{6}{\pi n}\sum_{i=1}^{n-1} l(t_i t_{i+1}^{-1}) + \frac{2}{\pi n} l(t_n t_{n-1}) \leq 6\lambda(t)
\end{aligned}$$

works and yields $\ell_1'(g) \leq 6\tilde{\lambda}(g)$.

In Sp$(2n)$ the characters are given by $\beta_i(t) = t_i t_{i+1}^{-1}$ for $i < n$ and by $\beta_n(t) = t_n^2$.

We estimate

$$\begin{aligned}
\lambda(t) &= \frac{1}{\pi n}\sum_{i=1}^{n-1} l(t_i t_{i+1}^{-1}) + \frac{1}{\pi n} l(t_n^2) \\
&\leq \frac{1}{2n}\sum_{i=1}^{n-1} |1 - z t_i (z t_{i+1})^{-1}| + \frac{1}{2n}|1 - z t_n z^{-1} t_n| \\
&\leq \frac{1}{2n}\sum_{i=1}^{n-1} |1 - z t_i| + |1 - z t_i^{-1}| + \frac{3}{4n}(|1 - z t_n| + |1 - z t_n^{-1}|) \\
&\leq \frac{3}{4n}\sum_{i=1}^{2n} |1 - z t_i| = \frac{3 \cdot 2n}{2n}\ell_1(zt).
\end{aligned}$$

This implies $\tilde{\lambda}(g) \leq \frac{3}{2}\ell_1'(g)$. We have the same possibilities to permute t_i, t_j and t_i, t_i^{-1} as in $\mathrm{SO}(2n+1)$. Therefore we can estimate

$$\begin{aligned}
\ell_1'(t) &= \inf_{z \in Z(\mathrm{U}(2n))} \frac{1}{2n}\sum_{i=1}^{n} |1 - t_i z| + |1 - t_i^{-1} z| \\
&\leq \inf_{z \in Z(\mathrm{U}(2n))} \frac{2}{\pi n}\sum_{i=1}^{n} l(t_i z) + l(t_i^{-1} z) \\
&\leq \frac{2}{\pi n}\sum_{i=1}^{n-1} l(t_i t_n) + l(t_i^{-1} t_n) + \frac{2}{\pi n} l(t_n^2) \\
&\leq \frac{4}{\pi n}\sum_{i=1}^{n-1} l(t_i t_{i+1}^{-1}) + \frac{2}{\pi n} l(t_n^2) \leq 4\lambda(t),
\end{aligned}$$

and finally $\ell_1'(g) \leq 4\tilde{\lambda}(g)$ for all $g \in \mathrm{Sp}(2n)$. □

We remark that, as is clear from the proof, the pair (L, L') in Lemma 6.4 can be chosen equal to $(2,2)$, $(6,1)$, $(6,2)$ or $(4,\frac{3}{2})$ to work for groups $\mathrm{SU}(n+1)$, $\mathrm{SO}(2n+1)$, $\mathrm{SO}(2n)$ or $\mathrm{Sp}(2n)$, respectively.

Note that in contrast to the treatment of finite simple groups we have to consider exceptional groups, since for Lie groups of bounded rank, instead of Theorem 5.12, we will be facing the more complicated Theorem 6.14 below.

LEMMA 6.5 *There is $L > 0$ such that the following holds: Let G be an exceptional compact connected quasisimple Lie group of rank r. Then for every $g \in G$ satisfying $\tilde{\lambda}(g) \leq \frac{1}{2r}$ the estimates*

$$L^{-1}\ell_1'(g) \leq \lambda(g) \leq L\ell_1'(g)$$

hold, where ℓ'_1 is defined as in Lemma 6.4.

Proof. Suppose the standard representation δ' embeds G into $\mathrm{U}(2n)$, and let $\omega_1, \ldots, \omega_n$ be the weights of δ. Then δ' has the weights $\omega_1, \ldots \omega_n$ and $\omega_{n+i} := \omega_i^{-1}$ for $i = 1 \ldots n$, and the diagonal elements in $\mathrm{U}(2n)$ coming from t in a maximal torus T of G take the form $\operatorname{diag}(\omega_1(t), \ldots, \omega_{2n}(t))$. Since the root lattice is contained in the weight lattice, every fundamental root is a linear combination of weights with integer coefficients, $\beta_i = \sum_{j=1}^n m_{ij}\omega_j$ say. Note that for an element $y \in S^1$ we have $|1-y| + |1-y^{-1}| \leq 2(|1-zy| + |1-zy^{-1}|)$ for any $z \in S^1$ if $l(y) \leq \frac{1}{2}\pi$. Let $m_i := \sum_{j=1}^n |m_{ij}|$. Since we assumed $\tilde{\lambda}(t) \leq \frac{1}{2r}$, $l(\beta_i(t)) \leq \frac{1}{2}\pi$ for every i, and we estimate

$$
\begin{aligned}
|1-\beta_i(t)| &= \tfrac{1}{2}(|1-\beta_i(t)| + |1-\beta_i(t)^{-1}|) \\
&\leq |1 - z^{m_i}\beta_i(t)| + |1 - z^{m_i}\beta_i(t)^{-1}| \\
&= \left|1 - \prod_{j=i}^{n} z^{|m_{ij}|}\omega_j(t)^{m_{ij}}\right| + \left|1 - \prod_{j=i}^{n} z^{|m_{ij}|}\omega_j(t)^{-m_{ij}}\right| \\
&\leq \sum_{j=1}^{n} |m_{ij}| \cdot |1 - z\omega_j(t)^{\operatorname{sgn} m_{ij}}| + \sum_{j=1}^{n} |m_{ij}| \cdot |1 - z\omega_j(t)^{-\operatorname{sgn} m_{ij}}| \\
&= \sum_{j=1}^{n} |m_{ij}|(|1 - z\omega_j(t)^{m_{ij}}| + |1 - z\omega_j(t)^{-m_{ij}}|).
\end{aligned}
$$

By summing over all $i = 1 \ldots r$ we obtain

$$
\sum_{i=1}^{r} |1-\beta_i(t)| \leq \sum_{i=1}^{r}\sum_{i=1}^{n} |m_{ij}| \cdot |1-\omega_j(t)^{\pm 1}| \leq \sum_{j=1}^{2n} M|1 - z\omega_j(t)|,
$$

where $M := \max_{j=1\ldots n} \sum_{i=1}^r |m_{ij}|$. By appropriate scaling of the two sums and taking the infimum over all z we arrive at $\lambda(t) \leq M\ell'_1(t)$ and also $\tilde{\lambda}(g) \leq M\ell'_1(g)$ for all $g \in G$.

By looking up the tables in [28], Appendix C, we find that the highest weight ω for δ is a linear combination of fundamental roots with integer coefficients. All other weights differ from ω by an element of the root lattice and therefore every weight is a linear combination of the fundamental roots with integer coefficients. We write $\omega_j = \sum_{i=1}^r n_{ji}\beta_i$, where $n_{ji} \in \mathbb{Z}$. Then, using the special number $z = 1$, similarly to the above calculation

$$\sum_{j=1}^{n} |1 - z\omega_j(t)^{\pm 1}| \leq \sum_{j=1}^{n} |1 - \omega_j(t)^{\pm 1}| \leq \tfrac{1}{2} \sum_{i=1}^{r} N|1 - \beta_i(t)^{\pm 1}|,$$

where $N := \max_{i=1\dots r} \sum_{j=1}^{n} |n_{ji}|$. After rescaling $\ell_1'(t) \leq 2N\tilde{\lambda}(t)$ follows. Setting $L := \max(M, 2N)$ finishes the proof. □

§ 3 BOUNDED GENERATION IN COMPACT CONNECTED SIMPLE LIE GROUPS

The motivation for this paragraph is taken from [34], Paragraph 5.5.4. The goal is to refine the methods from ibid. to obtain the result that in compact connected simple Lie groups an element which is not much longer, in the sense of measuring with λ, than some other element, can be written as a bounded product of conjugates of the latter.

Let G be a compact connected simple Lie groups with maximal torus T. Then every element $g \in T$ can be decomposed into the product of commuting factors $g = g_i \cdot g_i'$, where $g_i \in H_i$ and $g_i' \in T_i$. Moreover $\beta_i(g) = \beta_i(g_i)\beta_i(g_i') = \beta_i(g_i)$, because $g_i' \in T_i = \ker(\beta_i)$.

The following result can be found in the proof of Lemma 5.20 in [34].

LEMMA 6.6 *Let $G = \mathrm{SU}(2)$ and g, h be non-trivial elements in G such that $\lambda(g) \leq m\lambda(h)$, $m \geq 2$ an integer. Then g is a product of at most m conjugates of h.*

In the notation introduced in § 1 the image of φ_α in G is isomophic to SU(2) or PSU(2), depending on whether G is simply connected or not. Since we want to apply Lemma 6.6, it would be convenient to work in simply connected groups. Because the classical quasisimple Lie groups defined in § 1 are not necessarily simply connected, we have to use covering groups instead. Since the functions ℓ_1' and $\tilde{\lambda}$ are defined modulo the center, it will in the end make no difference which group on the scale from simply connected to simple we used.

We continue with further features of the internal structure of a (simply connected) compact connected quasisimple Lie group, as outlined in [34]. We make adjustments to the text and notation of this reference when needed.

We give two lemmas concerning linear combinations of roots and the resulting impact on the structure of G.

LEMMA 6.7 *In any simple root system Φ every long root β can be written as $\beta = \alpha_1 + \alpha_2$ for short roots α_1 and α_2. Every short root α can be written as $\alpha = \mu\beta_1 + \mu\beta_2$, where $\mu \in \pm\{\frac{1}{3}, \frac{1}{2}, 1\}$. These are the only coefficients that can appear in a linear combination of two roots to a third.*

P r o o f. The lemma follows from inspection of the standard representations of root systems.

If Φ is of type A_n, D_n, E_6, E_7 or E_8 all roots have the same length and there is nothing to prove. In case of Φ being of type B_n, the roots are exactly the integer vectors v in $\mathbb{R}^n$ with Euclidean norm $|v| = 1$ or $|v| = \sqrt{2}$. For type C_n we have $\Phi = \left\{v \in \mathbb{Z}^n \,\middle|\, |v| = \sqrt{2}\right\} \cup \{v \in (2\mathbb{Z})^n \,|\, |v| = 2\}$. We see that in these cases $\mu = \pm\frac{1}{2}$. If Φ is of type F_4, it is the union of the set of all vectors in $\mathbb{R}^4$ with two or one components equal to ± 1 and the others equal to 0 and the set of vectors with all components being $\pm\frac{1}{2}$. Here μ is either $\pm\frac{1}{2}$ or ± 1, depending on the short root. In the remaining case of type G_2 we represent Φ by vectors in $\mathbb{R}^3$, the short roots being

$$(1,-1,0), (-1,1,0), (1,0,-1), (-1,0,1), (0,1,-1), (0,-1,1)$$

and the long roots

$$(2,-1,-1), (-2,1,1), (1,-2,1), (-1,2,-1), (1,1,-2), (-1,-1,2).$$

Again a close look implies the claim with $\mu = \pm\frac{1}{3}$. □

LEMMA 6.8 *Let α, β be fundamental roots of different lengths and g an element in H_α such that $l(\alpha(g)) = \varepsilon$. Then there are elements w_1 and w_2 in the Weyl group such that $H_\alpha \subset H_\beta^{w_1} H_\beta^{w_2}$ and in particular g equals the product $g_1 g_2$, where $g_i \in H_\beta^{w_i}$ are elements such that $l(\beta^{w_i}(g_i)) \leq \varepsilon$.*

Proof. The inclusion $H_\alpha \subset H_\beta^{w_1} H_\beta^{w_2}$ can be found in [34] in the proof of Lemma 5.19 and the argument goes as follows. The Weyl group W acts on the roots. There are elements w_1 and w_2 in W such that α equals the linear combination $\mu_1\beta^{w_1} + \mu_2\beta^{w_2}$. The claim follows.

We have to go into detail and take care of lengths of elements in the product. To each root δ corresponds a coroot h_δ in the Lie algebra of G. Chapter 23 in [8] shows that there is a normalization of coroots such that we can assume $h_{\delta+\zeta} = h_\delta + h_\zeta$. The homomorphism η_δ is induced by $e_\delta : \vartheta \mapsto \exp(\vartheta i h_\delta)$. (Confer [39], Theorems 6.20, 4.8, 4.16.) Here the angle $l(\delta(e_\delta(\vartheta)))$ equals $2|\vartheta|$ if $|\vartheta| \in [0, \frac{1}{2}\pi]$ and $2\pi - 2|\vartheta|$ if $|\vartheta| \in [\frac{1}{2}\pi, \pi]$. From Lemma 6.7 we know that the coefficients $\mu := \mu_1 = \mu_2$ are $\pm\frac{1}{3}$, $\pm\frac{1}{2}$ or ± 1. Hence μ^{-1} is an integer and if we write $\gamma_i := \beta^{w_i}$, $i = 1, 2$,

$$h_{\gamma_1} + h_{\gamma_2} = h_{\gamma_1+\gamma_2} = h_{\mu^{-1}\alpha} = \mu^{-1} h_\alpha$$

and so the coroots obey the same linear relation as the roots.

Assume without loss of generality $g = e_\alpha(\frac{1}{2}\varepsilon)$, where $\varepsilon < \pi$. Then $l(\alpha(g)) = \varepsilon$ and $\alpha = \mu\gamma_1 + \mu\gamma_2$ implies

$$e_\alpha(\tfrac{1}{2}\varepsilon) = \exp(\tfrac{1}{2}\varepsilon i h_\alpha) = \exp(\tfrac{1}{2}\mu\varepsilon i h_{\gamma_1}) \exp(\tfrac{1}{2}\mu\varepsilon i h_{\gamma_2}) = e_{\gamma_1}(\tfrac{1}{2}\mu\varepsilon) e_{\gamma_2}(\tfrac{1}{2}\mu\varepsilon).$$

Hence $g = e_\alpha(\frac{1}{2}\varepsilon) \in H_\alpha$ is the product of elements $g_i = e_{\gamma_i}(\frac{1}{2}\mu\varepsilon)$ in the subgroups $H_\beta^{w_1}$ and $H_\beta^{w_2}$, respectively, with angle $l(\gamma_i(g_i)) = \mu\varepsilon \leq \varepsilon$. $\square$

We are now ready to generalize Lemma 6.6 to an arbitrary compact connected quasisimple Lie group G. We use the interplay of the groups S_i and the Weyl group W, and the decomposition of the maximal torus T into subgroups H_i. To use Lemma 6.6 we can in the following safely assume that G is a simply connected Lie group, since λ is zero on the center of G and thus well defined on the quotient $G/Z(G)$.

We use the abbreviation $h^{-H} := (h^{-1})^H$ for subgroups H of G and $h \in G$.

LEMMA 6.9 *Let $g_i \in H_i$ and $h_j \in H_j$, corresponding to g and h in T, such that $l(\beta_i(g)) \leq m\, l(\beta_j(h))$, where m is an even integer. Then*

$$g_i \in (h^G \cup h^{-G})^{4m}.$$

Proof. The proof splits in two cases whether β_i and β_j have the same length or not.

If β_i and β_j are roots of the same length, then there is an element v in the Weyl group W such that $H_i^v = H_j$. This entails $g_i^v \in H_j \subset S_j$. The Weyl group permutes the roots by conjugating the argument. Hence $l(\beta_j(g_i^v)) = l(\beta_i(g_i))$, and $g_i^v \in \left(h_j^{S_j}\right)^m$ by Lemma 6.6. We compute

$$(h^{S_j})^m = ((h_j h_j')^{S_j})^m = \left(h_j^{S_j}\right)^m \cdot h_j'^m$$

to deduce

$$g_i \in ((h^{S_j})^m h_j'^{-m})^{v^{-1}}.$$

Now $l(\beta_j(1)) = 0$ and by Lemma 6.6, $1 \in (h_j^{S_j})^2$. Therefore, and because h_j' commutes with every element in S_j,

$$(h_j')^2 \in \left(h_j^{S_j}\right)^2 \cdot (h_j')^2 = (h^{S_j})^2.$$

Note that this works equally well for h^{-1} instead of h. Because we assumed m even, we arrive at

$$g_i \in ((h^{S_j})^m (h^{-S_j})^m)^{v^{-1}} \subset (h^G \cup h^{-G})^{2m}.$$

If β_i and β_j are roots of different lengths, Lemma 6.8 gives the existence of elements w_1 and w_2 such that $g_i = f_1^{w_1} f_2^{w_2}$, where f_k is in H_j and $l(\beta_j(f_k)) \leq l(\beta_i(g_i))$, for $k = 1, 2$. Then, again by Lemma 6.6, $f_k \in \left(h_j^{S_j}\right)^m$, and we obtain

$$g_i \in \left(\left(h_j^{S_j}\right)^m\right)^{w_1} \cdot \left(\left(h_j^{S_j}\right)^m\right)^{w_2}.$$

We now proceed as above to deduce

$$g_i \in ((h^{S_j})^m (h^{-S_j})^m)^{w_1} \cdot ((h^{S_j})^m (h^{-S_j})^m)^{w_2} \subset (h^G \cup h^{-G})^{4m}. \qquad \square$$

We need a variant of the previous lemma for classical groups, involving several roots at once:

LEMMA 6.10 *Let G be a classical group and g and h be elements in T. Let β_{i_k}, β_{j_k} be roots of the same length, where $k = 1 \dots s$. Assume that β_{i_k} and β_{i_l} as well as β_{j_k} and β_{j_l} are pairwise orthogonal, i.e. $|i_k - i_l| \geq 2$ and $|j_k - j_l| \geq 2$ if $l \neq k$. If $\beta_{i_l}(g) \leq m\beta_{j_l}(h)$ for all $l = 1 \dots s$, where m is an even integer, then*

$$g_{i_1} \cdot g_{i_2} \cdot \dots \cdot g_{i_s} \in (h^G \cup h^{-G})^{2m}.$$

P r o o f. The proof is similar to the one of Lemma 6.9 and requires only one additional argument from [34]. We write $h = h_{j_1} \cdot \dots \cdot h_{j_s} \cdot \overline{h}$, where $\overline{h} \in T_{j_1} \cap \dots \cap T_{j_s}$, and similarly $g = g_{i_1} \cdot \dots \cdot g_{i_s} \cdot \overline{g}$. For every choice of distinct k, l the groups S_{j_k} and S_{j_l} commute elementwise, because the corresponding roots are orthogonal. Since $\overline{h} \in T_{j_1} \cap \dots \cap T_{j_s}$, also $\overline{h}$ commutes with S_{j_k} for all k. Hence

$$\left(h^{S_{j_1} \cdot \dots \cdot S_{j_s}}\right)^m = \left(h_{j_1}^{S_{j_1}} \cdot \dots \cdot h_{j_s}^{S_{j_s}} \cdot \overline{h}^{S_{j_1} \cdot \dots \cdot S_{j_s}}\right)^m = \left(h_{j_1}^{S_{j_1}}\right)^m \cdot \dots \cdot \left(h_{j_s}^{S_{j_s}}\right)^m \cdot \overline{h}^m$$

follows. By Lemma 5.21 in [34] and orthogonality of the roots involved there is an element v in the Weyl group of G such that $H_{i_k}^v = H_{j_k}$ for all $k = 1 \dots s$. By Lemma 6.6 $g_{i_k}^v \in \left(h_{j_k}^{S_{j_k}}\right)^m$ for all $k = 1 \dots s$. This implies

$$g_{i_1} \cdot \dots \cdot g_{i_s} \in \left(\left(h^{S_{j_1} \cdot \dots \cdot S_{j_s}}\right)^m \cdot \overline{h}^{-m}\right)^{v^{-1}}.$$

We also have

$$\overline{h}^{-2} \in \left(h_{j_1}^{S_{j_1}}\right)^{-2} \cdot \dots \cdot \left(h_{j_s}^{S_{j_s}}\right)^{-2} \cdot \overline{h}^{-2} = \left(h^{S_{j_1} \cdot \dots \cdot S_{j_s}}\right)^{-2},$$

which results in

$$g_{i_1} \cdot \dots \cdot g_{i_s} \in \left(\left(h^{S_{j_1} \cdot \dots \cdot S_{j_s}}\right)^m \cdot \left(h^{-S_{j_1} \cdot \dots \cdot S_{j_s}}\right)^m\right)^{v^{-1}} \subset (h^G \cup h^{-G})^{2m},$$

as claimed □

The next theorem is modelled after Case 1 in Lemma 5.19 in [34].

THEOREM 6.11 *Let $\varepsilon > 0$ and G be a compact connected simple Lie group of rank r. Assume g and h are non-trivial elements in T satisfying $\lambda(h) = \varepsilon$ and $\lambda(g) \leq m\lambda(h)$ for an even integer number m. Then*

$$g \in (h^G \cup h^{-G})^{4mr^2}.$$

P r o o f. Write $g = g_1 \cdot \ldots \cdot g_r$, $h = h_1 \cdot \ldots \cdot h_r$, where $g_i, h_i \in H_i$. For reasons of averaging there is one fundamental root β_j such that $l(\beta_j(h)) \geq \varepsilon\pi$. Let $m_i \geq 2$ be the smallest even integer such that $\lambda(g_i) \leq m_i \lambda(h_j)$. Then m_i cannot be larger than mr for any i. In the worst case, there might be many m_i which are very small (such as $m_i = 2$), while some m_i is close to mr. Therefore we must estimate $\sum_{i=1}^{r} m_i \leq mr^2$. We now use Lemma 6.9 to obtain for all i

$$g_i \in (h^G \cup h^{-G})^{4m_i},$$

independently of the length of roots involved. Because g is the product of the g_i, and summing the m_i gives at most mr^2, g is a product of $4mr^2$ or less conjugates of h and h^{-1}. □

§ 4 LIE GROUPS OF BOUNDED RANK

We want to show that an analogue of Theorem 5.13 holds for well behaved Lie groups.

For the next theorem we also write $\tilde{\lambda}$ for the function obtained as a pointwise ultralimit of the functions $\tilde{\lambda}$ in ultraproducts of Lie groups.

THEOREM 6.12 *Let G_i be compact connected quasisimple Lie groups. If $\boldsymbol{g} \in \boldsymbol{G} := \prod_{\mathrm{u}} G_i$ satisfies $\lambda(\boldsymbol{g}) > 0$, then $N(\boldsymbol{g}) = \boldsymbol{G}$. The set $\boldsymbol{N}$ of all $\boldsymbol{g}$ such that $\tilde{\lambda}(\boldsymbol{g}) = 0$ is a normal subgroup and $\boldsymbol{G}/\boldsymbol{N}$ is simple.*

P r o o f. We can assume $g_i \in T_i$, where T_i is a maximal torus of G_i. Then the first part of the theorem follows already from Lemma 5.19 in [34]. For groups of bounded rank we can alternatively use Theorem 6.11.

By Lemma 6.4 and Lemma 6.5 $\tilde{\lambda}(\boldsymbol{g}) = 0$ is equivalent to $\ell_1'(\boldsymbol{g}) = 0$. Because ℓ_1' is a pseudo length function, $\boldsymbol{N}$ is a normal subgroup. From the first part of the theorem we deduce that $\boldsymbol{G}/\boldsymbol{N}$ is simple. □

We define $\boldsymbol{g} \preceq \boldsymbol{h}$ for $\boldsymbol{g}, \boldsymbol{h} \in \boldsymbol{G} \setminus \{1\}$ as in Section 5, § 4, except that we use ℓ_1' as our length of choice. Then Lemma 5.14 immediately implies $\boldsymbol{g} \preceq \boldsymbol{h}$ whenever $\boldsymbol{g} \in N(\boldsymbol{h})$.

LEMMA 6.13 *Let $\boldsymbol{G}$ be an ultraproduct of compact connected simple Lie groups G_i of bounded rank and assume $\boldsymbol{g} \preceq \boldsymbol{h}$ for non-trivial elements $\boldsymbol{g}$ and $\boldsymbol{h}$ in $\boldsymbol{G}$. Then $\boldsymbol{g} \in N(\boldsymbol{h})$.*

Proof. The hypothesis assures $\lambda(g_i) \leq m\lambda(h_i)$ for almost all i and a suitable constant m. Following Theorem 6.11 we immediately obtain $g_i \in C(h_i^{\pm 1})^{4mr^2}$, where r is the bound on the rank of the groups G_i. Hence

$$\boldsymbol{g} \in C(\boldsymbol{h}^{\pm 1})^{4mr^2} \subset N(\boldsymbol{h}) \qquad \square$$

We are now ready to prove the analogue of Theorem 5.13 for Lie groups of bounded rank.

THEOREM 6.14 *Let G_i be compact connected simple Lie groups of bounded rank. Then the set $\mathfrak{N}$ of normal subgroups of $\boldsymbol{G} := \prod_{\mathrm{u}} G_i$ is linearly ordered by inclusion.*

Proof. Exactly as in the proof of Theorem 5.13 we show that the set $\mathfrak{N}_0$ of normal closures of elements of $\boldsymbol{G}$ is order isomorphic to a subset of $\boldsymbol{K}/\equiv$. This is the quotient of $\boldsymbol{K} := \prod_{\mathrm{u}}[0,1]$ by the equivalence relation $\equiv$, which defines $\boldsymbol{a}$ and $\boldsymbol{b}$ equivalent if

$$0 < \lim_{\mathrm{u}} \frac{a_i}{b_i} < \infty.$$

Because a maximal torus T in a Lie group of rank r is isomorphic to the standard torus $(S^1)^r$, it is clear that for any prescribed a in $[0,1]$ there is an element in T with length a. Hence $\mathfrak{N}_0$ is isomorphic to $\boldsymbol{K}/\equiv$. Now an application of Lemma 5.18 shows that also $\mathfrak{N}$ is linearly ordered. $\square$

§ 5 LIE GROUPS OF UNBOUNDED RANK

Unfortunately, unlike for finite simple groups, Theorem 6.14 turns out to be false if there is no bound on the rank. We illustrate this fact as follows.

Let $G_n := \mathrm{SU}(2n+1)$. We consider elements

$$g_n = \operatorname{diag}\left(e^{\frac{\mathrm{i}2\pi(n-1)}{n}}, e^{\frac{\mathrm{i}\pi}{n^2}}, e^{\frac{\mathrm{i}\pi}{n^2}}, \ldots, e^{\frac{\mathrm{i}\pi}{n^2}}, 1, \ldots, 1\right), \qquad h_n = \operatorname{diag}(-1,-1,1,\ldots,1)$$

in the maximal torus T of diagonal matrices, where n entries of g_n equal 1. To make the counterexample meaningful we have to pass to $\mathrm{PSU}(2n+1)$, or equivalently use pseudo length functions that vanish on $Z = Z(\mathrm{SU}(2n+1))$. If we assume that $\boldsymbol{g} \in N(\boldsymbol{h})$, then Lemma 5.14 implies the existence of a constant m such that $\inf_{z\in Z} \ell_{\mathrm{r}}(z g_n) \leq m \inf_{z\in Z} \ell_{\mathrm{r}}(z h_n)$ for infinitely many n. But the left hand side converges to $\frac{1}{2}$ and the right hand side equals $\frac{2m}{2n+1}$, which does not fit together well. On the other hand $\boldsymbol{h} \in N(\boldsymbol{g})$ would imply

$$\tilde{\lambda}(h_n) \leq 2\ell_1'(h_n) \leq 2m\ell_1'(g_n) \leq 4m\tilde{\lambda}(g_n),$$

for some (other) constant m and infinitely many n. After ordering the entries of g_n and h_n appropriately we evaluate $\tilde{\lambda}(g_n) \leq \frac{4}{n(2n+1)}$ and $\tilde{\lambda}(h_n) = \frac{4}{2n+1}$ to obtain a contradiction once again, and must conclude that neither $N(\boldsymbol{g}) \subset N(\boldsymbol{h})$ nor $N(\boldsymbol{h}) \subset N(\boldsymbol{g})$ holds. Therefore the set of normal subgroups in the ultraproduct of (projective) special unitary groups cannot be linearly ordered.

Despite this setback we try to see how far we can get. Let g be an element in a classical compact connected quasisimple Lie group of rank n with maximal torus T. In what follows we will call $t \in C(g) \cap T$ **optimal** if the following holds: For all $s \in C(g) \cap T$ we have $|1 - \beta_1(t)| \geq |1 - \beta_1(s)|$, and for all $k = 1 \ldots n-1$ the equations $|1 - \beta_i(t)| = |1 - \beta_i(s)|$ for $i = 1 \ldots k$ imply $|1 - \beta_{k+1}(t)| \geq |1 - \beta_{k+1}(s)|$. We define a function $F_g : \mathbb{N} \to [0,1]$ by

$$F_g(i) := \begin{cases} \frac{1}{2}|1 - \beta_{\sigma(i)}(t)|, & i \in [n], \\ 0, & i > n, \end{cases}$$

where t is optimal and σ is a permutation of $[n]$ such that $F_g(i) \geq F_g(i+1)$ results for all $i \geq 1$. Note that there is always an optimal t and for different optimal elements s and t the differences $|1 - \beta_i(t)|$ and $|1 - \beta_i(s)|$ are the same. Hence F_g is well defined.

Let the sequences of functions (F_j) and (H_j) be representatives of elements $\boldsymbol{F}$ and $\boldsymbol{H}$, respectively, in the ultraproduct

$$\boldsymbol{M} := \prod_{\mathfrak{u}} \mathscr{F}_{n_j}(\mathbb{N}, [0,1]),$$

where $\mathscr{F}_{n_j}(\mathbb{N}, [0,1])$ is the set of decreasing functions $\mathbb{N} \to [0,1]$ with support contained in $[n_j]$. We let $\boldsymbol{F} \preceq \boldsymbol{H}$ if and only if there are constants c and $k \in \mathbb{N}$

such that for $\mathfrak{u}$-almost all j

$$F_j(ki+1) \leq cH_j(i+1),$$

whenever $i \geq 0$. It is clear that this defines a quasiorder on the space $\boldsymbol{M}$. We let $\boldsymbol{F} \equiv \boldsymbol{H}$ if $\boldsymbol{F} \preceq \boldsymbol{H}$ and $\boldsymbol{H} \preceq \boldsymbol{F}$ to obtain the the quotient space $\boldsymbol{M}/\equiv$ with the induced ordering. If $\boldsymbol{g} \in \boldsymbol{G} \setminus \{1\}$ we define $F_{\boldsymbol{g}}$ as the element in $\boldsymbol{M}$ associated with $(F_{g_j})_j$. With these notions at hand let $\boldsymbol{g} \preceq \boldsymbol{h}$ be equivalent to $F_{\boldsymbol{g}} \preceq F_{\boldsymbol{h}}$.

LEMMA 6.15 *Let g and h be elements in a classical compact connected quasisimple Lie group G. Then, if $i, j \geq 0$,*

$$F_{gh}(6i+6j+1) \leq 2F_g(i+1) + 2F_h(j+1).$$

P r o o f. Consider the standard embedding of G in $\mathrm{U}(n)$. The singular values $s_i(1-g)$ of $1-g$ are defined as

$$s_i(1-g) = |\lambda_i(1-g)|,$$

where $\lambda_i = \lambda_i(1-g)$ is some eigenvalue of $1-g$ counted with algebraic multiplicity, and we assume the $s_i(1-g)$ in decreasing order. We define

$$\underline{s}_i(g) := \inf_{z \in \mathrm{Z}(\mathrm{U}(n))} \tfrac{1}{2} s_i(1-zg).$$

Observe that since $s_{i+1}(1-zg) \leq s_i(1-zg)$ holds for all z, also $\underline{s}_{i+1}(g) \leq \underline{s}_i(g)$ is true for $i = 1 \dots n$. We prove the claim by showing two inequalities for the different groups in question.

Let first $G = \mathrm{SU}(n+1)$ and $t \in C(g) \cap T$ be optimal. Then $F_g(i) = \frac{1}{2}|t_{\sigma(i)} - t_{\sigma(i)+1}|$. Let $z \in \mathrm{Z}(\mathrm{U}(n+1))$ and τ a permutation such that

$$|z - t_{\tau(1)}| \geq |z - t_{\tau(2)}| \geq \ldots \geq |z - t_{\tau(n)}|.$$

If we assume the existence of i such that $|t_{\sigma(2i+1)} - t_{\sigma(2i+1)+1}| > 2|z - t_{\tau(i+1)}|$, then for all $k = 1 \dots 2i+1$ the estimate $2|z - t_{\tau(i+1)}| < |z - t_{\sigma(k)}| + |z - t_{\sigma(k)+1}|$ follows. Hence $|z - t_{\sigma(k)}| > |z - t_{\tau(i+1)}|$ or $|z - t_{\sigma(k)+1}| > |z - t_{\tau(i+1)}|$ and $\sigma(k) \in \{\tau(1), \dots, \tau(i)\}$ or $\sigma(k)+1 \in \{\tau(1), \dots, \tau(i)\}$. It might happen that $\sigma(k) = \sigma(l)+1$ for some $1 \leq k, l \leq i$, but not more than i times. Thus $\{\tau(1), \dots, \tau(i)\}$

contains at least $i+1$ elements, a contradiction. Therefore $F_g(2i+1) \le s_{i+1}(1-zg)$ holds for all i, independently of z.

If $i,j \ge 0$ and $i+j+1 \le n+1$ we have for central elements x,y in $\mathrm{U}(n+1)$ by the Ky Fan singular value inequality

$$\begin{aligned} s_{i+j+1}(1-xygh) &= s_{i+j+1}((1-xg)yh+(1-yh)) \\ &\le s_{i+1}((1-xg)yh)+s_{j+1}(1-yh) \\ &= s_{i+1}(1-xg)+s_{j+1}(1-yh). \end{aligned}$$

(Confer [19].) Combining this estimate with the previous one we obtain

$$F_{gh}(2i+2j+1) \le s_{i+1}(1-xg)+s_{j+1}(1-yh).$$

Taking the infimum over all $x,y \in Z(\mathrm{U}(n+1))$ on both sides yields

$$F_{gh}(2i+2j+1) \le 2\underline{s}_{i+1}(g)+2\underline{s}_{j+1}(h).$$

For the other classical groups the proof is similar and we will point out where slight changes have to be made. Consider the case $G=\mathrm{SO}(2n+1)\subset \mathrm{U}(2n+1)$. We assume that $|1-\beta_{\sigma(3i+1)}(t)| > 2|z-t_{\tau(i+1)}|$ (when $3i+1 \le n$), where τ satisfies $|z-t_{\tau(1)}| \ge |z-t_{\tau(2)}| \ge \ldots \ge |z-t_{\tau(2n+1)}|$. Since we have $\beta_n(t)=t_n$, for i_0 such that $\sigma(i_0)=n$ we must distinguish two cases. If $i_0=1$ the maximality of $|z-t_{\tau(1)}|$ is contradicted by

$$2|z-t_{\tau(1)}| < |1-\beta_{\sigma(1)}| \le |z-t_{\sigma(1)}|+|z-1|.$$

(Note that $|z-1|$ is a singular value of $z-g$.) For other i_0 the proof works as above, and we use the factor 3 to be able to ignore i_0 when deducing the contradiction. We then obtain $F_g(3i+1) \le s_{i+1}(1-zg)$ for all i, and with the help of the Ky Fan singular value inequality as above

$$F_{gh}(3i+3j+1) \le 2\underline{s}_{i+1}(g)+2\underline{s}_{j+1}(h)$$

follows.

In $\mathrm{SO}(2n)$ or $\mathrm{Sp}(2n)$ we do nearly the same as for $\mathrm{SO}(2n+1)$. We assume $|1-\beta_{\sigma(3i+1)}(t)| > 2|z-t_{\tau(i+1)}|$. For i_0 such that $\sigma(i_0)=n$ we obtain in $\mathrm{SO}(2n)$ a contradiction by the estimate

$$2|z - t_{\tau(1)}| < |1 - \beta_{\sigma(1)}| = |zz^{-1} - t_n t_{n-1}| \leq |z - t_n| + |z - t_{n-1}^{-1}|,$$

and in $\mathrm{Sp}(2n)$ similarly by $2|z - t_{\tau(1)}| < |z - t_n| + |z - t_n^{-1}|$. Since we chose the factor 3, we can again disregard the different root when arguing as in the case of $\mathrm{SU}(n+1)$ for other i_0. Hence for these groups as well

$$F_{gh}(3i + 3j + 1) \leq 2\underline{s}_{i+1}(g) + 2\underline{s}_{j+1}(h)$$

is true.

Now we proceed by showing a second estimate. Consider $\mathrm{SU}(n+1)$ and let t be optimal. Because we can conjugate with arbitrary generalized permutation matrices of determinant 1 in $\mathrm{SU}(n+1)$, by definition of optimality $|t_i - t_{i+1}| \geq |t_i - t_j|$ for all $i = 1 \ldots n-1$, $j \geq i+1$. Let τ be a permutation such that $|t_{n-1} - t_{\tau(1)}| \geq |t_{n-1} - t_{\tau(2)}| \geq \ldots \geq |t_{n-1} - t_{\tau(n)}|$. Then

$$\begin{aligned} 2\underline{s}_i(g) &= \inf_{z \in Z(\mathrm{U}(n))} s_i(z - t) \leq s_i(t_n - t) \\ &= |t_n - t_{\tau(i)}| \leq |t_{\tau(i)} - t_{\tau(i)+1}| \leq 2F_g(\sigma^{-1}\tau(i)) \end{aligned}$$

if $\tau(i) \leq n-1$. If $\tau(i) = n$ we have $\underline{s}_i(g) \leq 0$. Because F_g is decreasing by definition and we can estimate each $\underline{s}_i(g)$ from above with a unique value of F_g, in fact

$$\underline{s}_i(g) \leq F_g(i)$$

follows for all $i = 1 \ldots n$.

Now let $G = \mathrm{SO}(2n+1)$ and t be optimal. Then $|t_i - t_{i+1}| \geq |t_i - t_j^{\pm 1}|$ for all $i = 1 \ldots n-1$, $i+1 \leq j \leq n$. Let τ be a permutation such that $|t_n - t_{\tau(1)}| \geq |t_n - t_{\tau(2)}| \geq \ldots \geq |t_n - t_{\tau(n)}|$. We have to take into account that $|t_n - t_i|$ and $|t_n - t_i^{-1}|$ might be of comparable size. Therefore, using $\underline{s}_{2i+1}(g)$ instead of $\underline{s}_{1+i}(g)$, for $i \geq 0$

$$\begin{aligned} 2\underline{s}_{2i+1}(g) &\leq s_{2i+1}(t_n - t) \leq |t_n - t_{\tau(i+1)}| \\ &\leq |t_{\tau(i+1)} - t_{\tau(i+1)+1}| \leq 2F_g(\sigma^{-1}\tau(i+1)) \end{aligned}$$

if $\tau(i+1) \neq n$, and $\underline{s}_{2i+1}(g) \leq 0$ otherwise. Hence $\underline{s}_{2i+1}(g) \leq F_g(i)$ for all $i = 0 \ldots n-1$.

If $G = \mathrm{SO}(2n)$ we have $|t_i - t_{i+1}| \geq |t_i - t_j^{\pm 1}|$ for all $i = 1 \ldots n-2$, $i+1 \leq j \leq n$. Let τ be a permutation as above corresponding to t_{n-1}. Then we proceed as before to obtain $\underline{s}_{2i+1}(g) \leq F_g(i)$ for all $i = 0 \ldots n-1$. The case of $\mathrm{Sp}(2n)$ is similar.

Combining the different estimates finishes the proof. □

PROPOSITION 6.16 *Let $\boldsymbol{g}$ and $\boldsymbol{h}$, both not equal to 1, be elements in an ultraproduct of compact connected simple Lie groups G_j such that $\boldsymbol{g} \in N(\boldsymbol{h})$. Then $\boldsymbol{g} \preceq \boldsymbol{h}$.*

Proof. Let G_j have rank r_j and consider the interesting case $r_j \to_{\mathfrak{u}} \infty$. We assume that $\boldsymbol{g}$ is a product of k conjugates of $\boldsymbol{h}^{\pm 1}$. This implies that $g_j \in G_j$ is a product of not more than $6k$ conjugates of $h_j^{\pm 1}$ for $\mathfrak{u}$-almost all j. By conjugating we can assume g_j and h_j in a maximal torus of G_j. We only have to take care of r_j sufficiently larger than $6k$. Imagine G_j embedded in a unitary group by the standard representation, in order to use Lemma 6.15. In a group of such a large rank now for $i \geq 0$

$$F_{g_j}(6ki+1) \leq 2^k 6k F_{h_j}(i+1)$$

holds, because F_{h_j} is invariant under conjugation of h_j with unitaries. □

A graph X has **coloring number** $\chi(X) = k$ if there is a coloring of the vertices with k colors such that no two vertices of the same color are joined by an edge and k is minimal with this property.

Let X be a graph with a partition of the vertices into subsets of size k. Then a **strong k-coloring** of X is a coloring such that every color appears in each partition exactly once. (If the number of vertices is not divisible by k we add isolated vertices as needed.) Then the **strong coloring number** $s\chi(X)$ of X is the least k such that for all partitions of the vertices into subsets of size k, X admits a strong k-coloring. (Confer also [2].)

LEMMA 6.17 *There is a natural number $s \geq 3$ such that the following holds. Let σ be a permutation of the numbers $[n]$, where n is divisible by s. Then one* *can partition* $(\sigma(1),\sigma(2),\ldots,\sigma(n))$ *into* s *vectors* $v_i := (a_{i,1},\ldots,a_{i,\frac{n}{s}})$ *such that* $|a_{i,j}-a_{i,k}| \neq 1$ *for all* $i = 1\ldots s$ *and* $j,k = 1\ldots\frac{n}{s}$, *and* $a_{i,j} = \sigma(k)$ *implies* $|sj-k| \leq s-1$.

P r o o f. We reformulate the problem in graph theoretical terms. Consider the vector $(1,2,\ldots,n)$ as a graph, where i,j are connected if $|i-j| \in \{1,n-1\}$, i.e. the cycle C_n. We assign a vertex i of this graph the label $\lfloor\frac{\sigma^{-1}(i)+s-1}{s}\rfloor$. Thus we use $\frac{n}{s}$ different labels, each one occuring exactly s times. If $s \geq s\chi(C_n)$, then there is a proper coloring of C_n such that no two vertices with the same label have the same color. Now let v_i be the vector of vertices of color i, in the ordering prescribed by the labels. Then it follows immediately that no two consecutive numbers appear in the same v_i. If $a_{i,j} = \sigma(k)$, then $j = \lfloor\frac{k+s-1}{s}\rfloor$ and the difference $|sj-k|$ is strictly less than s.

Since it is known that the strong coloring number of C_n can be bounded independently of n, the claim follows. □

Note that the constant $s\chi(C_n)$ in the previous lemma can be made explicit. Alon in [2] mentions the bound of $s\chi(C_n) \leq 4$ (for n divisible by 4), credited to de la Vega, Fellows and himself. The usual proofs invoke probabilistic methods such as the Lovász local lemma. Fleischner and Stiebnitz proved in [20] that $s\chi(C_n) = 3$, and there is an elementary proof for this fact, presented by Sachs in [38].

LEMMA 6.18 *Let* G *be a compact connected simple Lie group of rank* $r > 20k$ *for a natural number* k. *Assume* g *and* h *are non-trivial elements in the maximal torus* T *satisfying* $F_g(ki+1) \leq mF_h(i+1)$ *if* $i \geq 0$, *where* $m \in \mathbb{N}$ *is an even integer. Then*

$$g \in (h^G \cup h^{-G})^{140km+4m}.$$

P r o o f. Considering the rank requirements, G is a classical group of type A_r, B_r, C_r or D_r. Then without loss of generality the roots β_i, $i = 1\ldots r-1$, form a root system of type A_{r-1} and the root β_r possibly has a different length. Roots β_j and β_i are orthogonal, whenever $|i-j| \geq 2$ and we will say that i is orthogonal to j in that case.

We can assume without loss of generality that g and h are optimal. Let $K := 5k$. With N the largest natural number divisible by 3 such that $NK \leq r - K - 1$, we define N-tuples $A_l := (l, K+l, 2K+l, \ldots, NK+l)$ for $l = 1 \ldots K$ and $A_0 := (1, 2, \ldots, N)$. We choose the permutation σ implicitly by writing $F_g(i) = \frac{1}{2}|1 - \beta_{\sigma(i)}(g)|$ as above. Likewise we have τ corresponding to h. Both permutations act coordinatewise on N-tuples. If we choose $i \geq 0$, then

$$\tfrac{1}{2}|1-\beta_{\sigma(Ki+l)}(g)| = F_g(Ki+l) \leq mF_h(i+1) = \tfrac{1}{2}m|1-\beta_{\tau(i+1)}(h)|.$$

If $l \in \{1, \ldots, K\}$ we hence obtain

$$l(\beta_{\sigma(Ki+l)}(g)) \leq 2ml(\beta_{\tau(i+1)}(h)).$$

Without loss of generality, we can assume the worst case that $\tau(A_0)$ contains N consecutive numbers. Then by Lemma 6.17 and the subsequent remarks there is a partition of $\tau(A_0)$ into tuples B_1, B_2 and B_3 with the same number of elements such that the entries in B_i are pairwise orthogonal for $i = 1, 2, 3$. In the same way we obtain $C_{i,l}$, $i = 1, 2, 3$, from the sets $\sigma(A_l)$, $l = 1 \ldots K$. By Lemma 5.21 in [34] (here we use orthogonality) there are elements $w_{i,l}$ in the Weyl group of G that map the vectors of fundamental roots corresponding to B_i to the vectors of fundamental roots corresponding to $C_{i,l}$ for all $i = 1, 2, 3$, $l = 1 \ldots k$. We will apply Lemma 6.10 to all g_i for $i \in C_{j,l}$, but have to check first if we will end up with good enough constants. Lemma 6.17 with $s = 3$ guarantees that indices are at a distance of at most 2 from their optimal position. In the worst case we have to compare $l(\beta_{\sigma(K(i-2)+l)}(g))$ with $l(\beta_{\tau(i+3)}(h))$. Since under this assumption $i \geq 3$,

$$\begin{aligned} K(i-2)+l &\geq 5k(i-2)+1 = ki+1+4ki-10k \\ &\geq (ki+1)+12k-10k = k(i+2)+1. \end{aligned}$$

This implies

$$l(\beta_{\sigma(K(i-2)+l)}(g)) \leq l(\beta_{\sigma(k(i+2)+1)}(g)) \leq 2ml(\beta_{\tau(i+3)}(h)).$$

For other possibly dislocated indices this kind of estimate works as well. After the abovementioned application of Lemma 6.10 we know

$$\prod_{i \in C_{j,l}} g_i \in (h^G \cup h^{-G})^{2 \cdot 2m}.$$

When we reconstruct most of g in this way, we arrive at

$$\prod_{i\in\bigcup C_{j,l}} g_i \in (h^G \cup h^{-G})^{4m\cdot 15k},$$

because we had to treat $3\cdot K = 15k$ sets $C_{j,l}$. What remains are the indices left out in the above procedure. The number of these is $r - NK \leq 4K$ by the choice of N. If $i \leq r-1$, using $h_{\tau(1)}$, we can generate the g_i separately in $2\cdot 2m$ steps as in Lemma 6.9. The last root β_r possibly requires the second argument in the proof of Lemma 6.9, which results in adding $4\cdot 2m$. Hence generating the missing parts of g can be done in $(4K-1)\cdot 4m + 8m = 4(20k+1)m$ steps.

All in all we end up with

$$g \in (h^G \cup h^{-G})^{140km+4m}$$

as claimed. □

THEOREM 6.19 *Let $\boldsymbol{g}$ and $\boldsymbol{h}$ be elements in the ultraproduct $\boldsymbol{G}$ of compact connected simple Lie groups of unbounded rank. Then $\boldsymbol{g} \preceq \boldsymbol{h}$ is equivalent to $\boldsymbol{g} \in N(\boldsymbol{h})$.*

Proof. The first implication was already proved in Proposition 6.16. The proof of the second is an application of Lemma 6.18, analogous to the proofs of Theorem 5.13 or Theorem 6.14. □

Up to now it is clear that the set of normal closures of elements in $\boldsymbol{G}$ is order isomorphic to $\boldsymbol{M}/\equiv$. What remains to be clarified is the influence of this ordering on the ordering of normal subgroups.

We are interested in the lattice of normal subgroups of groups G. The lattice operations are $N \wedge M = N \cap M$ and $N \vee M = NM$, the normal subgroup generated by N and M, for any choice of normal subgroups in G. It is well known that the lattice of normal subgroups of any group is modular, that is for normal subgroups L, M and N the modular law

$$((L\wedge N)\vee M)\wedge N = (L\wedge N)\vee(M\wedge N)$$

holds.

LEMMA 6.20 *Let $\boldsymbol{G}$ be an ultraproduct of compact connected simple Lie groups G_j and $\boldsymbol{g}$, $\boldsymbol{h}$ in the ultraproduct $\boldsymbol{T} \subset \boldsymbol{G}$ of maximal tori of the G_j. Then there are $\boldsymbol{a}, \boldsymbol{b} \in \boldsymbol{T}$ such that $N(\boldsymbol{g}) \wedge N(\boldsymbol{h}) = N(\boldsymbol{a})$ and $N(\boldsymbol{g}) \vee N(\boldsymbol{h}) = N(\boldsymbol{b})$.*

P r o o f. We define functions $\boldsymbol{A} := \min(F_g, F_h)$ and $\boldsymbol{B} := \max(F_g, F_h)$. The plan is to show that there are actually elements $\boldsymbol{a}$ and $\boldsymbol{b}$ such that $\boldsymbol{A} = F_a$ and $\boldsymbol{B} = F_b$. For some j consider the functions $A_j := \min(F_{g_j}, F_{h_j})$, $B_j := \min(k \mapsto 1, \max(F_{g_j}, F_{h_j}))$. Let T_j be a maximal torus in the group G_j of rank r, where we can assume $g_j, h_j \in T_j$. Because T_j is isomorphic to $(S^1)^r$ we find elements a_j and b_j in T_j such that $F_{a_j} = A_j$ and $F_{b_j} = B_j$. This yields $\boldsymbol{a}$ and $\boldsymbol{b}$ as claimed. □

PROPOSITION 6.21 *Let $\boldsymbol{G}$ be an ultraproduct of compact connected simple Lie groups. Then the set $\mathfrak{N}_0$ of normal closures of elements in $\boldsymbol{G} \setminus \{1\}$ is a distributive lattice.*

P r o o f. We already know that $\mathfrak{N}_0$ is order isomorphic to $\boldsymbol{M}/\equiv$. It is clear that the latter is a distributive lattice with meet and join induced by the operations min and max applied to functions. Lemma 6.20 shows that the corresponding operations in $\mathfrak{N}_0$ produce normal closures again. □

LEMMA 6.22 *Let G be a group. If the set of normal closures of elements in G is a distributive lattice, then the lattice of normal subgroups is distributive, too.*

P r o o f. Let L, M and N be any normal subgroups in G. We have to show that

$$(L \vee M) \wedge N = (L \wedge N) \vee (M \wedge N)$$

holds. Here the inclusion of the right hand side in the left hand side is true in general. Moreover by assumption the whole equation holds for normal closures of elements in G. Consider $x \in (L \vee M) \wedge N$. Then $x \in L \vee M$ and $x \in N$ because the meet operation is intersection of sets. Because the normal closure of L and M is the normal subgroup LM, there are $a \in L$ and $b \in M$ such that x equals the product ab. This means that $x \in N(a) \vee N(b)$. We also observe $N(x) \subset N$ to obtain

$$\begin{aligned} x &\in (N(a) \vee N(b)) \wedge N(x) \\ &= (N(a) \wedge N(x)) \vee (N(b) \wedge N(x)) \\ &\subset (L \wedge N) \vee (M \wedge N). \end{aligned}$$

Thus the claim follows. □

The observations made in Proposition 6.21 and Lemma 6.20 suffice to prove the following result.

THEOREM 6.23 *If G is an ultraproduct of compact connected simple Lie groups, then the lattice of normal subgroups of G is distributive.*

§ 6 THE LATTICE OF NORMAL SUBGROUPS IN ULTRAPRODUCTS OF COMPACT SIMPLE GROUPS

The Peter-Weyl Theorem ([8], Theorem 4.2) implies that every compact topological group G has a unitary representention, i.e. a continuous homomorphism φ into $\mathrm{U}_n(\mathbb{C})$. By continuity the image $\varphi(G)$ is a compact group again. If Γ is simple, then φ is injective and thus G isomorphic to a compact subgroup of a Lie group. By Cartan's Theorem G is a Lie group again. Since the component of the identity is a normal subgroup, if G is not connected it is discrete, in which case it is finite by compactness. Thus the compact simple groups are exactly the compact connected simple Lie groups and finite simple groups.

We considered ultraproducts of finite simple groups and compact connected simple Lie groups. We have to deal with the subcases of groups of bounded and unbounded rank, because the two behave differently as shown above. If we have an ultraproduct $\boldsymbol{G}$ of compact simple groups the ultrafilter selects one kind of groups among the four listed possibilities which determine the properties of $\boldsymbol{G}$. We will say that $\boldsymbol{G}$ is of **bounded finite type**, **unbounded finite type**, **bounded Lie type** or **unbounded Lie type** if $\boldsymbol{G}$ is essentially an ultraproduct of finite simple groups of bounded or unbounded rank or Lie groups of bounded or unbounded rank, respectively.

Recall the situation in the case of finite simple groups. We defined $\boldsymbol{g} \preceq \boldsymbol{h}$ if

$$\lim_{u} \frac{\ell(g_i)}{\ell(h_i)} < \infty,$$

where ℓ was one of the length functions ℓ_{H} and ℓ_{J}. For $g \neq 1$ in a finite simple group of rank n (where A_n has rank n) define

$$F_g(k) := \begin{cases} 0, & \text{otherwise,} \\ 1, & \text{if } k \le n\ell(g). \end{cases}$$

Then it is an elementary observation that $F_g \preceq F_h$ if and only if $\boldsymbol{g} \preceq \boldsymbol{h}$ for non-trivial $\boldsymbol{g}, \boldsymbol{h} \in \boldsymbol{G}$, where the comparison of functions is in the same sense as for Lie groups above. Using this last remark we can summarize our results in the following theorem.

THEOREM 6.24 *Let $\boldsymbol{G}$ be an ultraproduct of non-abelian compact simple groups G_j. Let $\boldsymbol{M}$ be the ultraproduct of sequences of decreasing functions $F_j : \mathbb{N} \to [0,1]$ with support of size less or equal to the rank of G_j. Define $\boldsymbol{F} \preceq \boldsymbol{H}$ if there are constants c, k such that $F_j(ki+1) \le cH_j(i+1)$ for all $i \ge 0$ $\mathfrak{u}$-almost everywhere, and $\boldsymbol{F} \equiv \boldsymbol{H}$ if $\boldsymbol{F} \preceq \boldsymbol{H}$ as well as $\boldsymbol{H} \preceq \boldsymbol{F}$.*

(1) *If $\boldsymbol{G}$ is of unbounded Lie type, then the set of normal closures $\mathfrak{N}_0$ of elements in $\boldsymbol{G} \setminus \{1\}$ is a lattice isomorphic to the distributive lattice $\boldsymbol{M}/\equiv$. The lattice $\mathfrak{N}$ of normal subgroups of $\boldsymbol{G}$ is distributive.*

(2) *If $\boldsymbol{G}$ is of bounded Lie type, then $\mathfrak{N}_0$ is isomorphic to the linearly ordered sublattice of $\boldsymbol{M}/\equiv$ induced by the functions of bounded support and $\mathfrak{N}$ is linearly ordered.*

(3) *If $\boldsymbol{G}$ is of unbounded finite type, then $\mathfrak{N}_0$ is isomorphic to the linearly ordered sublattice of $\boldsymbol{M}/\equiv$ induced by the functions $F : \mathbb{N} \to \{0,1\}$. Again, $\mathfrak{N}$ is linearly ordered.*

(4) *If $\boldsymbol{G}$ is of bounded finite type, then $\boldsymbol{G}$ is simple and $\mathfrak{N}$ is isomorphic to the lattice 2.*

Index of Definitions

REFERENCES

[1] John Allsup and Richard Kaye, *Normal subgroups of nonstandard symmetric and alternating groups,* Arch. Math. Logic **46** (2007), no. 2, 107–121.

[2] Noga Alon, *The strong chromatic number of a graph,* Random Structures Algorithms **3** (1992), no. 1, 1–7.

[3] Goulnara Arzhantseva and Liviu Păunescu, *Linear sofic groups and algebras,* available at http://arxiv.org/abs/1212.6780.

[4] Itaï Ben Yaacov, Alexander Berenstein, C. Ward Henson and Alexander Usvyatsov, *Model theory for metric structures,* in Model theory with applications to algebra and analysis. Vol. 2, volume 350 of London Math. Soc. Lecture Note Ser., pages 315–427. Cambridge Univ. Press, Cambridge (2008).

[5] Armand Borel, *Linear algebraic groups,* volume 126 of Graduate Texts in Mathematics, Springer-Verlag, New York, second edition (1991).

[6] Nicolas Bourbaki, *Elements of mathematics. General topology. Part 1,* Hermann, Paris (1966).

[7] Nathanial P. Brown and Narutaka Ozawa, *C^*-algebras and finite-dimensional approximations,* volume 88 of Graduate Studies in Mathematics. American Mathematical Society, Providence, RI (2008).

[8] Daniel Bump, *Lie groups*, volume 225 of Graduate Texts in Mathematics. Springer-Verlag, New York, (2004).

[9] Stanley N. Burris and Hanamantagouda P. Sankappanavar, *A course in universal algebra,* Millenium edition (2000).

[10] Roger W. Carter, *Simple groups of Lie type,* Wiley Classics Library. John Wiley & Sons Inc., New York (1989). Reprint of the 1972 original, A Wiley-Interscience Publication.

[11] Benoît Collins and Kenneth J. Dykema, *Free products of sofic groups with amalgamation over monotileably amenable groups,* Münster J. Math. **4** (2011), 101–117.

[12] Gertrude Ehrlich, *Characterization of a continuous geometry within the unit group,* Trans. Amer. Math. Soc. **83** (1956), 397–416.

[13] David Eisenbud, *Commutative algebra with a view toward algebraic geometry,* volume 150 of Graduate Texts in Mathematics. Springer-Verlag, New York, (1995).

[14] Gábor Elek and Endre Szabó, *Sofic groups and direct finiteness,* J. Algebra **280** (2004), no. 2, 426–434.

[15] Gábor Elek and Endre Szabó, *Hyperlinearity, essentially free actions and L^2-invariants. The sofic property,* Math. Ann. **332** (2005), no. 2, 421–441.

[16] Gábor Elek and Endre Szabó, *On sofic groups,* J. Group Theory **9** (2006), no. 2, 161–171.

[17] Gábor Elek and Endre Szabó, *Sofic representations of amenable groups,* Proc. Amer. Math. Soc. **139** (2011), no. 12, 4285–4291.

[18] Paul Ellis, Sherwood Hachtman, Scott Schneider and Simon Thomas, *Ultraproducts of finite alternating groups,* RIMS Kokyu-roku (2008), no. 1619, 1–7.

[19] Ky Fan, *Maximum properties and inequalities for the eigenvalues of completely continuous operators,* Proc. Nat. Acad. Sci., U. S. A. **37** (1951), 760–766.

[20] Herbert Fleischner and Michael Stiebitz, *A solution to a colouring problem of P. Erdős,* Discrete Math. (1992), no. 101, 39–48.

[21] Felix R. Gantmacher, *The theory of matrices. Vol. 1.* AMS Chelsea Publishing, Providence, RI (1998). Translated from the Russian by K. A. Hirsch, Reprint of the 1959 translation.

[22] Lev Glebsky and Luis Manuel Rivera, *Sofic groups and profinite topology on free groups,* J. Algebra **320** (2008), no. 9, 3512–3518.

[23] Mikhael Gromov, *Endomorphisms of symbolic algebraic varieties,* J. Eur. Math. Soc. (JEMS) **1** (1999), no. 2, 109–197.

[24] Larry C. Grove, *Classical groups and geometric algebra,* volume 39 of Graduate Studies in Mathematics. American Mathematical Society, Providence, RI (2002).

[25] Karl W. Gruenberg, *Residual properties of infinite soluble groups,* Proc. London Math. Soc. (3) **7** (1957), 29–62.

[26] Graham Higman, *A finitely generated infinite simple group,* J. London Math. Soc. **26** (1951), 61–64.

[27] Kei-ichiro Iima and Ryo Iwamatsu, *On the Jordan decomposition of tensored matrices of Jordan canonical forms,* Math. J. Okayama Univ. **51** (2009), 133–148.

[28] Anthony W. Knapp, *Lie groups beyond an introduction,* volume 140 of Progress in Mathematics. Birkhäuser Boston Inc., Boston, MA, second edition (2002).

[29] Serge Lang, *Introduction to algebraic geometry,* Addison-Wesley Publishing Co., Inc., Reading, Mass. (1972). Third printing, with corrections.

[30] Serge Lang, *Algebra,* volume 211 of Graduate Texts in Mathematics. Springer-Verlag, New York, third edition (2002).

[31] Martin W. Liebeck and Aner Shalev, *Diameters of finite simple groups: sharp bounds and applications,* Ann. of Math. (2) **154** (2001), no. 2, 383–406.

[32] Fumitomo Maeda, *Kontinuierliche Geometrien,* Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete, Bd. 95. Springer-Verlag, Berlin (1958).

[33] Nikolay Nikolov, *Strange images of profinite groups,* available at http://arxiv.org/abs/0901.0244.

[34] Nikolay Nikolov and Dan Segal, *Generators and commutators in finite groups; abstract quotients of compact groups,* Invent. Math. **190** (2012), no. 3, 513–602.

[35] Liviu Păunescu, *On sofic actions and equivalence relations,* J. Funct. Anal. **261** (2011), no. 9, 2461–2485.

[36] Vladimir G. Pestov, *Dynamics of infinite-dimensional groups,* volume 40 of University Lecture Series. American Mathematical Society, Providence, RI (2006). The Ramsey-Dvoretzky-Milman phenomenon, Revised edition of ıt Dynamics of infinite-dimensional groups and Ramsey-type phenomena [Inst. Mat. Pura. Apl. (IMPA), Rio de Janeiro (2005); MR2164572].

[37] Vladimir G. Pestov, *Hyperlinear and sofic groups: a brief guide,* Bull. Symbolic Logic **14** (2008), no. 4, 449–480.

[38] Hans Sachs, *Elementary proof of the cycle-plus-triangles theorem,* in Combinatorics, Paul Erdős is eighty, Vol. 1, Bolyai Soc. Math. Stud., pages 347–359. János Bolyai Math. Soc., Budapest (1993).

[39] Mark R. Sepanski, *Compact Lie groups,* volume 235 of Graduate Texts in Mathematics. Springer, New York (2007).

[40] Abel Stolz and Andreas Thom, *On the lattice of normal subgroups in ultraproducts of compact simple groups,* available at http://arxiv.org/abs/1207.0977.

[41] Andreas Thom, *About the metric approximation of Higman's group,* J. Group Theory **15** (2012), no. 2, 301–310.

[42] John von Neumann, *Continuous geometry,* Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ (1998). With a foreword by Israel Halperin, Reprint of the 1960 original, Princeton Paperbacks.

[43] Benjamin Weiss, *Sofic groups and dynamical systems*, Sankhyā Ser. A **62** (2000), no. 3, 350–359. Ergodic theory and harmonic analysis. Mumbai (1999).

[44] Robert A. Wilson, *The finite simple groups*, volume 251 of Graduate Texts in Mathematics. Springer-Verlag London Ltd., London (2009).

Selbständigkeitserklärung

Hiermit erkläre ich, die vorliegende Dissertation selbständig und ohne unzulässige fremde Hilfe angefertigt zu haben. Ich habe keine anderen als die angeführten Quellen und Hilfsmittel benutzt und sämtliche Textstellen, die wörtlich oder sinngemäß aus veröffentlichten oder unveröffentlichten Schriften entnommen wurden, und alle Angaben, die auf mündlichen Auskünften beruhen, als solche kenntlich gemacht. Ebenfalls sind alle von anderen Personen bereitgestellten Materialen oder erbrachten Dienstleistungen als solche gekennzeichnet.

Leipzig, den 23. Oktober 2013

Abel Stolz